

CODEN: JASMAN

The Journal of the Acoustical Society of America

ISSN: 0001-4966

Vol. 117, No. 4

April 2005

ACOUSTICAL NEWS—USA		1675
USA Meetings Calendar		1675
ACOUSTICAL NEWS—INTERNATIONAL		1679
International Meetings Calendar		1679
ABSTRACTS FROM ACOUSTICS RESEARCH LETTERS ONLINE		1681
BOOK REVIEWS		1683
OBITUARIES		1685
REVIEWS OF ACOUSTICAL PATENTS		1687
<hr/>		
LETTERS TO THE EDITOR		
An artificial neural network approach for predicting architectural speech security (L)	Jingfeng Xu, John S. Bradley, Bradford N. Gover	1709
Estimates of auditory filter phase response at and below characteristic frequency (L)	Andrew J. Oxenham, Stephan D. Ewert	1713
GENERAL LINEAR ACOUSTICS [20]		
Design and optimization of a noise reduction system for infrasonic measurements using elements with low acoustic impedance	Benoit Alcoverro, Alexis Le Pichon	1717
Analysis of the sound field in finite length infinite baffled cylindrical ducts with vibrating walls of finite impedance	Wei Shao, Chris K. Mechefske	1728
Acoustical measurement of the shear modulus for thin porous layers	Jean F. Allard, Michel Henry, Laurens Boeckx, Philippe Leclaire, Walter Lauriks	1737
Computation of scattering from clusters of spheres using the fast multipole method	Nail A. Gumerov, Ramani Duraiswami	1744
Modeling of phased array transducers	Rais Ahmad, Tribikram Kundu, Dominique Placko	1762
Lamb mode conversion at edges. A hybrid boundary element–finite-element solution	José M. Galán, Ramón Abascal	1777
Dispersion of waves in porous cylinders with patchy saturation: Formulation and torsional waves	James G. Berryman, Steven R. Pride	1785
A first-order statistical smoothing approximation for the coherent wave field in random porous random media	Tobias M. Müller, Boris Gurevich	1796

(Continued)

CONTENTS—Continued from preceding page

NONLINEAR ACOUSTICS [25]

Linear and nonlinear propagation of higher order modes in hard-walled circular ducts containing a real gas	Stefan Scheichl	1806
Simultaneous measurement of acoustic and streaming velocities in a standing wave using laser Doppler anemometry	Michael W. Thompson, Anthony A. Atchley	1828
Influences of a temperature gradient and fluid inertia on acoustic streaming in a standing wave	Michael W. Thompson, Anthony A. Atchley, Michael J. Maccarone	1839
On the interaction of counterpropagating acoustic waves in resonant rods composed of materials with hysteretic quadratic nonlinearity	Vitalyi Gusev	1850
On the stability of the effective apodization of the nonlinearly generated second harmonic with respect to range	Russell J. Fedewa, Kirk D. Wallace, Mark R. Holland, James R. Jago, Gary C. Ng, Brent S. Robinson, Matthew R. Rielly, James G. Miller	1858

AEROACOUSTICS, ATMOSPHERIC SOUND [28]

Acoustic pulse propagation through a fluctuating stably stratified atmospheric boundary layer	Igor Chunchuzov, Sergey Kulichkov, Alexander Otrezov, Vitaly Perepelkin	1868
The use of impedance matching capillaries for reducing resonance in rosette infrasonic spatial filters	Michael A. H. Hedlin, Benoit Alcoverro	1880
Simultaneous acoustic channel measurement via maximal-length-related sequences	Ning Xiang, John N. Daigle, Mendel Kleiner	1889

UNDERWATER SOUND [30]

Phase speed attenuation in bubbly liquids inferred from impedance measurements near the individual bubble resonance frequency	Preston S. Wilson, Ronald A. Roy, William M. Carey	1895
A phase-space Gaussian beam summation representation of rough surface scattering	Goren Gordon, Ehud Heyman, Reuven Mazar	1911
Phase-space beam summation analysis of rough surface waveguide	Goren Gordon, Ehud Heyman, Reuven Mazar	1922
Geoacoustic inversion in time domain using ship of opportunity noise recorded on a horizontal towed array	Cheolsoo Park, Woojae Seong, Peter Gerstoft	1933
Effects of environmental uncertainties on sonar detection performance prediction	Liewei Sha, Loren W. Nolte	1942
Bayesian sonar detection performance prediction in the presence of interference in uncertain environments	Liewei Sha, Loren W. Nolte	1954
Concurrent inversion of geo- and bio-acoustic parameters from transmission loss measurements in the Yellow Sea	Orest Diachok, Stephen Wales	1965
Long range acoustic imaging of the continental shelf environment: The Acoustic Clutter Reconnaissance Experiment 2001	Purnima Ratilal, Yisan Lai, Deanelle T. Symonds, Lilimar A. Ruhlmann, John R. Preston, Edward K. Scheer, Michael T. Garr, Charles W. Holland, John A. Goff, Nicholas C. Makris	1977
Time-reversal imaging for classification of submerged elastic targets via Gibbs sampling and the Relevance Vector Machine	Nilanjan Dasgupta, Lawrence Carin	1999
Protocols for calibrating multibeam sonar	Kenneth G. Foote, Dezhang Chu, Terence R. Hammar, Kenneth C. Baldwin, Larry A. Mayer, Lawrence C. Hufnagle, Jr., J. Michael Jech	2013

CONTENTS—Continued from preceding page

ULTRASONICS, QUANTUM ACOUSTICS, AND PHYSICAL EFFECTS OF SOUND [35]

- | | | |
|---|-----------------------------------|------|
| Multi-mode Lamb wave tomography with arrival time sorting | Kevin R. Leonard, Mark K. Hinders | 2028 |
|---|-----------------------------------|------|

STRUCTURAL ACOUSTICS AND VIBRATION [40]

- | | | |
|---|---|------|
| Regularization method for measurement of structural intensity using nearfield acoustical holography | Kenji Saijyou, Chiaki Okawara | 2039 |
| Smart panels with velocity feedback control systems using triangularly shaped strain actuators | Paolo Gardonio, Stephen J. Elliott | 2046 |
| Reconstruction of transient acoustic radiation from a sphere | Sean F. Wu, Huancai Lu, Manjit S. Bajwa | 2065 |

NOISE: ITS EFFECTS AND CONTROL [50]

- | | | |
|--|---|------|
| Analytical approach for sound attenuation in perforated dissipative silencers with inlet/outlet extensions | A. Selamet, M. B. Xu, I.-J. Lee, N. T. Huff | 2078 |
|--|---|------|

ARCHITECTURAL ACOUSTICS [55]

- | | | |
|--|--------------------------------|------|
| Absorptive properties of rigid porous media: Application to face centered cubic sphere packing | S. Gasser, F. Paun, Y. Bréchet | 2090 |
|--|--------------------------------|------|

ACOUSTIC SIGNAL PROCESSING [60]

- | | | |
|--|--|------|
| Theory and design of sound field reproduction in reverberant rooms | Terence Betlehem, Thushara D. Abhayapala | 2100 |
| Microphone array signal processing with application in three-dimensional spatial hearing | Mingsian R. Bai, Chenpang Lin | 2112 |
| The shift-invariant discrete wavelet transform and application to speech waveform analysis | Jörg Enders, Weihua Geng, Peijun Li, Michael W. Frazier, David J. Scholl | 2122 |
| Defect imaging with guided waves in a pipe | Takahiro Hayashi, Morimasa Murase | 2134 |

PHYSIOLOGICAL ACOUSTICS [64]

- | | | |
|--|---|------|
| Simplified nonlinear outer hair cell models | Niranjan Deo, Karl Grosh | 2141 |
| Modeling high-frequency electromotility of cochlear outer hair cell in microchamber experiment | Zhijie Liao, Aleksander S. Popel, William E. Brownell, Alexander A. Spector | 2147 |

PSYCHOLOGICAL ACOUSTICS [66]

- | | | |
|--|--|------|
| Modulation masking produced by second-order modulators | Christian Füllgrabe, Brian C. J. Moore, Laurent Demany, Stephan D. Ewert, Stanley Sheft, Christian Lorenzi | 2158 |
| The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners | Tanya L. Arbogast, Christine R. Mason, Gerald Kidd, Jr. | 2169 |

SPEECH PERCEPTION [71]

- | | | |
|---|---|------|
| A Speech Intelligibility Index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners | Koenraad S. Rhebergen, Niek J. Versfeld | 2181 |
| Perception of pitch location within a speaker's F0 range | Douglas N. Honorof, D. H. Whalen | 2193 |
| Perception of aperiodicity in pathological voice | Jody Kreiman, Bruce R. Gerratt | 2201 |
| Consonant recognition and the articulation index | Jont B. Allen | 2212 |

(Continued)

CONTENTS—Continued from preceding page

Coherence and the speech intelligibility index	James M. Kates, Kathryn H. Arehart	2224
SPEECH PROCESSING AND COMMUNICATION SYSTEMS [72]		
Statistical properties of infant-directed versus adult-directed speech: Insights from speech recognition	Katrin Kirchhoff, Steven Schimmel	2238
MUSIC AND MUSICAL INSTRUMENTS [75]		
The acoustics of Japanese wooden drums called “mokugyo”	Masahiro Sunohara, Kenji Furihata, David K. Asano, Takesaburo Yanagisawa, Atsuyoshi Yuasa	2247
Beating frequency and amplitude modulation of the piano tone due to coupling of tones	Bo Cartling	2259
Generation of longitudinal vibrations in piano strings: From physics to sound synthesis	Balázs Bank, László Sujbert	2268
Aerodynamic excitation and sound production of blown-closed free reeds without acoustic coupling: The example of the accordion reed	Denis Ricot, René Caussé, Nicolas Misdariis	2279
BIOACOUSTICS [80]		
Measurement of the depth-dependent resonance of water-loaded human lungs	J. S. Martin, P. H. Rogers, E. A. Cudahy	2291
Instrumenting free-swimming dolphins echolocating in open water	Stephen W. Martin, Michael Phillips, Eric J. Bauer, Patrick W. Moore, Dorian S. Houser	2301
Echolocation characteristics of free-swimming bottlenose dolphins during object detection and identification	Dorian Houser, Stephen W. Martin, Eric J. Bauer, Michael Phillips, Tim Herrin, Matt Cross, Andrea Vidal, Patrick W. Moore	2308
Quantifying the acoustic repertoire of a population: The vocalizations of free-ranging bottlenose dolphins in Fiordland, New Zealand	Oliver Boisseau	2318
Information theory analysis of patterns of modulation in the advertisement call of the male bullfrog, <i>Rana catesbeiana</i>	Dianne N. Suggs, Andrea Megela Simmons	2330
Gas bubble and solid sphere motion in elastic media in response to acoustic radiation force	Yurii A. Ilinskii, G. Douglas Meegan, Evgenia A. Zabolotskaya, Stanislav Y. Emelianov	2338
The mechanical and thermal effects of focused ultrasound in a model biological material	Feng Feng, Ajit Mal, Michael Kabo, Jeffrey C. Wang, Yoseph Bar-Cohen	2347
CUMULATIVE AUTHOR INDEX		2356

ACOUSTICAL NEWS—USA

Elaine Moran

Acoustical Society of America, Suite 1N01, 2 Huntington Quadrangle, Melville, NY 11747-4502

Editor's Note: Readers of this Journal are encouraged to submit news items on awards, appointments, and other activities about themselves or their colleagues. Deadline dates for news items and notices are 2 months prior to publication.

The 148th meeting of the Acoustical Society of America held in San Diego, California

The 148th meeting of the Acoustical Society of America was held 15–19 November at the Town and Country Hotel in San Diego, California. This is the seventh time the Society has met in this city, the previous times being 1952, 1969, 1976, 1983, 1990, 1997.

The meeting drew a total of 1111 registrants, including 195 nonmembers and 230 students. Attesting to the international ties of our organization, 140 of the registrants, (that is, 13%) were from outside North America. There were 23 registrants from France, 22 from Japan, 19 from Germany, 17 from the United Kingdom, 10 from Korea, 6 from Italy, 4 each from Australia, Israel, Russia and Taiwan, 3 each from Denmark, New Zealand, People's Republic of China, and Portugal, 2 each from Brazil, Ireland, Norway and Sweden, and 1 each from Austria, Belgium, Finland, Malaysia, Spain, The Netherlands and Turkey. North American countries, the United States and Canada, accounted for 946 and 25 respectively.

A total of 770 papers, organized into 79 sessions, covered the areas of interest of all 13 Technical Committees. The meeting also included 7 meetings dealing with standards. The Monday evening tutorial lecture series was continued by Gerald D' Spain, Scripps Institution of Oceanography and Doug Wartok of Florida International University. Their tutorial, "Ocean Noise and Marine Mammals" was presented to an audience of about 135.

The Short Course on Acoustical Oceanography in Shallow Water was held on Sunday and Monday and the course instructors were George Frisk, Florida International University, Grand Deane, Scripps Institution of Oceanography, James Preisig, Woods Hole Oceanographic Institution and Dajun Tang, University of Washington. Fourteen registrants attended the short course.

The Society's thirteen Technical Committees held open meetings during the San Diego meeting where they made plans for special sessions at upcoming ASA meetings, discussed topics of interest to the attendees and held informal socials after the end of the official business. These meetings are working, collegial meetings and all people attending Society meetings are encouraged to attend and to participate in the discussions. More information about Technical Committees, including minutes of meetings, can be found on the ASA Website <<http://asa.aip.org/committees.html>>.

The equipment exhibit drew 13 exhibitors, including ACO Pacific, Inc., Bruel & Kjaer, Casella CEL USA, Eckel Industries, Acoustic Div., G.R.A.S. Sound and Vibration APS, Head Acoustics, Inc., MBI Productions Company, Inc., Navon Engineering Network, PCB Piezotronics/Larson Davis, Precision Filters Inc., Quest Technologies, Reson, and Springer Verlag New York. Exhibits included microphones, sound level alarms and monitoring systems, analyzers, sound absorbing materials and noise control enclosures, acoustic acquisition recorders, transducers, calibrators, baffles, software, accelerometers, sound level meters, dosimeters underwater acoustic transducers, and books on acoustics. An exhibit opening reception was held on Monday evening.

The ASA Student Council organized a Fellowship Workshop to provide information to help students and post-doctoral researchers obtain funding. Representatives from several government agencies attended including: National Science Foundation (NSF), Naval Research Laboratory (NRL), National Institutes of Health/National Institute of Deafness and Other Communication Disorders (NIH/NIDCD), the Office of Naval Research (ONR) and the Acoustical Society of America Prizes and Special Fellowships committee. The workshop drew over 70 attendees including undergraduates, graduate students, post-docs and faculty. After the workshop, the Council hosted a Student Reception with over 150 people in attendance. The Student Reception included the presentation of the inaugural Student Council Mentoring Award to David Blackstock (see Fig. 1).



FIG. 1. David Blackstock, recipient of the Student Council Mentoring Award.

Over 40 people participated in a technical tour and demonstration to visit two unique pipe organs in San Diego including the Spreckels Outdoor Pipe Organ in Balboa Park and a seven rank pipe organ at the First Methodist Church. Lyle Blackinton, Pipe Organ Builder, was the tour leader.

Social events included the two social hours held on Tuesday and Thursday, a reception for students, a Fellows Lounge and a Fellows Luncheon and the morning coffee breaks. A special program for students to meet one-on-one with members of the ASA over lunch, which is held at each meeting, was organized by the Committee on Education in Acoustics.

Walter Munk, ASA Honorary Fellow, of the Scripps Institution of Oceanography, was the speaker at the Fellows Luncheon which was attended by over 120 people.

These social events provided the settings for participants to meet in relaxed settings to encourage social exchange and informal discussions. The Women in Acoustics Luncheon was held on Wednesday afternoon and was attended by over 90 people.

The Acoustical Society Foundation sponsored a dinner on Wednesday evening. The theme of the dinner was "Honoring Women in Acoustics and



FIG. 2. ASA President William Kuperman (l) presents the 2003 Science Writing Award for Journalists to Ian Sample (r).



FIG. 3. ASA President William Kuperman congratulates Steven C. Thompson, recipient of the 2003 Science Writing Award in Acoustics for Professionals in Acoustics.

in particular the contributions of Robert and Evelyn Young to the Acoustical Society of America.” The speaker was Julie Shimer, CEO of Vocera Communications, who presented “Acoustic Challenges of the Next Generation Communicator.

The plenary session included a business meeting of the Society, announcements, acknowledgment of the members and other volunteers who organized the meeting and the presentation of awards and certificates to newly-elected Fellows.

The 2003 Science Writing Award in Acoustics for Journalists was presented to Ian Sample for his article “The Sound of Sunshine” published in the *Guardian* newspaper on July 24, 2003 (see Fig. 2). The Science Writing Award for Professionals in Acoustics was awarded to Stephen C. Thompson for his article “Tutorial on Microphone Technologies for Directional Hearing Aids,” published in *The Hearing Journal* (see Fig. 3).

The first Rossing Prize in Acoustics Education was presented to Allan D. Pierce of Boston University Afor

The Silver Medal in Acoustical Oceanography was presented to D. Vance Holliday of BAE Systems, “for contributions to the study of marine life, from plankton to whales.” (see Fig. 5). The Silver Medal in Biomedical Ultrasound/Bioresponse to Vibration was presented to James G. Miller of Washington University “for contributions to ultrasonic tissue characterization and quantitative echocardiography.” (see Fig. 6). The Silver Medal in Engineering Acoustics was presented to John V. Bouyoucos of Hydroacoustics, Inc., “for the invention and development of hydraulically powered acoustic amplifiers for underwater use.” (see Fig. 7). Honorary Fellowship was conferred upon Walter H. Munk of the Scripps Institution of Oceanography “for the invention of ocean acoustic tomography.” (see Fig. 8).

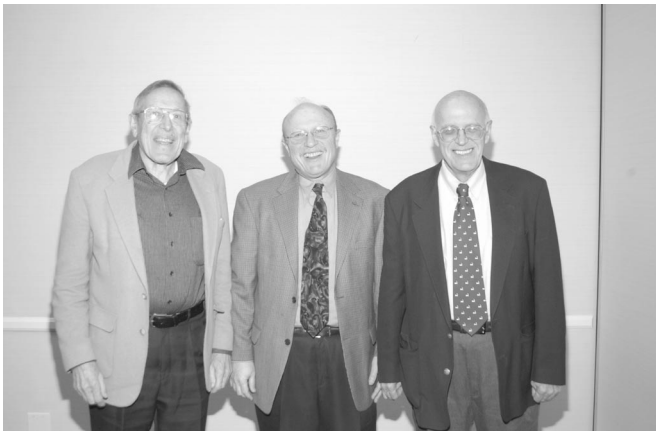


FIG. 4. Allan D. Pierce (r), recipient of the Rossing Prize in Acoustics Education receives congratulations from Thomas Rossing (l) and ASA President William Kuperman (c)



FIG. 5. ASA President William Kuperman presents Silver Medal in Acoustical Oceanography to D. Vance Holliday.

Election of 13 members to Fellow grade was announced and fellowship certificates were presented. New fellows are: Michael A. Ainslie, James J. Finneran, Klaus Genuit, Robert F. Gragg, K. Anthony Hoover, Patricia A. Keating, Volker Mellert, Robert E. Remez, Philippe Roux, Shihab Shamma, Yoiti Suzuki, Stephen C. Thompson and Manel E. Zakharia. (see Fig. 9)

ASA President William A. Kuperman expressed the Society’s thanks to the Local Committee for the excellent execution of the meeting, which clearly evidenced meticulous planning. Bill Kuperman also served as the Chair of the San Diego meeting (see Fig. 10) and he acknowledged the contributions of the members of his committee including: Michael J. Buckingham, Technical Program Chair, Patricia Jordan and Deborah Duckworth, Food Service/Social Events, Matthew A. Dzieciuch and Paul Baxley, Audio-Visual, Gail Smith, Accompanying Persons Program and Aaron Thode, Poster Sessions. He also expressed thanks to the members of the Technical Program Organizing Committee: Michael J. Buckingham, Technical Program Chair, Peter Gerstoft, Peter F. Worcester, Acoustical Oceanography; Ann E. Bowles, David A. Kastak, Animal Bioacoustics; Robin S. Glosemeyer, Architectural Acoustics; Francesco P. Curra, Biomedical Ultrasound/Bioresponse to Vibration; James P. Cottingham, Education in Acoustics and Musical Acoustics; Steven R. Baker, Engineering Acoustics; Alan H. Marsh, Noise; Robert M. Keolian, David L. Gardner, Physical Acoustics; Laura Dreisbach, Psychological and Physiological Acoustics; James V. Candy, Hassan H. Namarvar, Alireza A. Dibazar, Signal Processing in Acoustics; Dani Byrd, Jody Kreiman, Speech Communication; Courtney B. Burroughs, Structural Acoustics and Vibration; Stan E. Dosso, Kevin B. Smith, Underwater Acoustics; Melania Guerra, TPOM assistant.

The full technical program and award encomiums can be found in the



FIG. 6. ASA President William Kuperman congratulates James G. Miller, recipient of the Silver Medal in Biomedical Ultrasound/Bioresponse to Vibration



FIG. 7. ASA President William Kuperman congratulates John V. Bouyocos, recipient of the Silver Medal in Engineering Acoustics.

printed meeting program for readers who wish to obtain further information about the San Diego meeting. We hope that you will consider attending a future meeting of the Society to participate in the many interesting technical events and to meet with colleagues in both technical and social settings. Information on future meetings can be found in the *Journal* and on the ASA Home Page at <<http://asa.aip.org>>.

WILLIAM A. KUPERMAN
President 2004–2005

USA Meetings Calendar

Listed below is a summary of meetings related to acoustics to be held in the U.S. in the near future. The month/year notation refers to the issue in which a complete meeting announcement appeared.

2005	
16–20 May	149th Meeting joint with the Canadian Acoustical Association, Vancouver, Canada [Acoustical Society of America, Suite 1N01, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2360; Fax: 516-576-2377; Email: asa@aip.org ; WWW: http://asa.aip.org].
16–19 May	Society of Automotive Engineering Noise & Vibration Conference, Traverse City, MI [Patti Kreh, SAE International, 755 W. Big Beaver Rd., Ste. 1600, Troy, MI 48084, Tel.: 248-273-2474; Email: pkreh@sae.org].
18–22 July	17th International Symposium on Nonlinear Acoustics,



FIG. 8. ASA President William Kuperman congratulates Walter H. Munk, who was named Honorary Fellow of the Society.



FIG. 9. New Fellows of the Acoustical Society of America receive certificates from ASA President and Vice President.

State College, PA [Anthony Atchley, The Pennsylvania State University, 217 Applied Research Lab Building, University Park PA 16802; Tel.: 814-865-6364E-mail: ISNA17@outreach.psu.edu; WWW: <http://www.outreach.psu.edu/c&i/isna17/>

17–21 October
150th Meeting joint with Noise-Con 2005, Minneapolis, Minnesota, [Acoustical Society of America, Suite 1N01, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2360; Fax: 516-576-2377; Email: asa@aip.org; WWW: <http://asa.aip.org>].

Cumulative Indexes to the Journal of the Acoustical Society of America

Ordering information: Orders must be paid by check or money order in U.S. funds drawn on a U.S. bank or by Mastercard, Visa, or American Express credit cards. Send orders to Circulation and Fulfillment Division, American Institute of Physics, Suite 1N01, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2270. Non-U.S. orders add \$11 per index.

Some indexes are out of print as noted below.

Volumes 1–10, 1929–1938: JASA, and Contemporary Literature, 1937–1939. Classified by subject and indexed by author. Pp. 131. Price: ASA members \$5; Nonmembers \$10

Volumes 11–20, 1939–1948: JASA, Contemporary Literature and Patents. Classified by subject and indexed by author and inventor. Pp. 395. Out of Print

Volumes 21–30, 1949–1958: JASA, Contemporary Literature and Patents. Classified by subject and indexed by author and inventor. Pp. 952. Price:



FIG. 10. William Kuperman, Chair of the San Diego meeting

ASA members \$20; Nonmembers \$75

Volumes 31–35, 1959–1963: JASA, Contemporary Literature and Patents. Classified by subject and indexed by author and inventor. Pp. 1140. Price: ASA members \$20; Nonmembers \$90

Volumes 36–44, 1964–1968: JASA and Patents. Classified by subject and indexed by author and inventor. Pp. 485. Out of Print.

Volumes 36–44, 1964–1968: Contemporary Literature. Classified by subject and indexed by author. Pp. 1060. Out of Print

Volumes 45–54, 1969–1973: JASA and Patents. Classified by subject and indexed by author and inventor. Pp. 540. Price: \$20 (paperbound); ASA members \$25 (clothbound); Nonmembers \$60 (clothbound)

Volumes 55–64, 1974–1978: JASA and Patents. Classified by subject and indexed by author and inventor. Pp. 816. Price: \$20 (paperbound); ASA members \$25 (clothbound); Nonmembers \$60 (clothbound)

Volumes 65–74, 1979–1983: JASA and Patents. Classified by subject and

indexed by author and inventor. Pp. 624. Price: ASA members \$25 (paperbound); Nonmembers \$75 (clothbound)

Volumes 75–84, 1984–1988: JASA and Patents. Classified by subject and indexed by author and inventor. Pp. 625. Price: ASA members \$30 (paperbound); Nonmembers \$80 (clothbound)

Volumes 85–94, 1989–1993: JASA and Patents. Classified by subject and indexed by author and inventor. Pp. 736. Price: ASA members \$30 (paperbound); Nonmembers \$80 (clothbound)

Volumes 95–104, 1994–1998: JASA and Patents. Classified by subject and indexed by author and inventor. Pp. 632. Price: ASA members \$40 (paperbound); Nonmembers \$90 (clothbound)

Volumes 105–114, 1999–2003: JASA and Patents. Classified by subject and indexed by author and inventor. Pp.616 , Price: ASA members \$50; Nonmembers \$90 (paperbound)

ACOUSTICAL NEWS—INTERNATIONAL

Walter G. Mayer

Physics Department, Georgetown University, Washington, DC 20057

International Meetings Calendar

Below are announcements of meetings and conferences to be held abroad. Entries preceded by an * are new or updated listings.

April 2005

18–21 **International Conference on Emerging Technologies of Noise and Vibration Analysis and Control**, Saint Raphaël, France (Fax: +33 4 72 43 87 12; e-mail: goran.pavic@insa-lyon.fr)

May 2005

16–20 **149th Meeting of the Acoustical Society of America**, Vancouver, British Columbia, Canada (ASA, Suite 1N01, 2 Huntington Quadrangle, Melville, NY 11747-4502 USA; Fax: +1 516 576 2377; Web: asa.aip.org)

June 2005

1–3 ***1st International Symposium on Advanced Technology of Vibration and Sound**, Hiroshima, Japan (Web: dezima.ike.tottori-u.ac.jp/vstech2005)

20–23 **IEEE Oceans05 Europe**, Brest, France (ENST Bretagne—Technopôle Brest Iroise, 29238 Brest Cedex, France; Fax: +33 229 00 1098; Web: www.oceans05europe.org)

23–24 **2nd Congress of the Alps-Adria Acoustical Association (AAAA2005)**, Opatija, Croatia (Web: had.zea.fer.hr)

28–1 **International Conference on Underwater Acoustic Measurements: Technologies and Results**, Heraklion, Crete, Greece (Web: UAmesurements2005.iacm.forth.gr)

July 2005

4–8 **Turkish International Conference on Acoustics 2005: New Concepts for Harbor Protection, Littoral Security, and Underwater Acoustic Communications**, Istanbul, Turkey (Web: www.tica05.org/tica05)

11–14 ***12th International Congress on Sound and Vibration**, Lisbon, Portugal (Web: www.icsv12.ist.utl.pt)

August 2005

6–10 **Inter-Noise**, Rio de Janeiro, Brazil (Web: www.internoise2005.ufsc.br)

28–2 **EAA Forum Acusticum Budapest 2005**, Budapest, Hungary (I. Bába, OPAKFI, Fö u. 68, Budapest 1027, Hungary; Fax: +36 1 202 0452; Web: www.fa2005.org)

28–1 ***World Congress on Ultrasonics Merged with Ultrasonic International (WCU/UI'05)**, Beijing, China (Secretariat of WCU 2005, Institute of Acoustics, Chinese Academy of Sciences, P.O. Box 2712 Beijing, 100080 China; Fax: +86 10 62553898; Web: www.ioa.ac.cn/wcu-ui-05)

September 2005

4–8 **9th Eurospeech Conference (EUROSPEECH'2005)**, Lisbon, Portugal (Fax: +351 213145843; Web: www.interspeech2005.org)

5–9 **Boundary Influences in High Frequency, Shallow Water Acoustics**, Bath, UK (Web: acoustics2005.ac.uk)

18–21 ***IEEE International Ultrasonics Symposium**, Rotterdam, The Netherlands (Web: www.ieee-uffc.org)

20–22

***International Symposium on Environmental Vibrations**, Okayama, Japan (Web: isev2005.civil.okayama-u.ac.jp)

27–29

Autumn Meeting of the Acoustical Society of Japan, Sendai, Japan (Acoustical Society of Japan, Nakaura 5th-Bldg., 2-18-20 Sotokanda, Chiyoda-ku, Tokyo 101-0021, Japan; Fax: +81 3 5256 1022; Web: www.asj.gr.jp/index-en.html)

October 2005

12–14

***Acoustics Week in Canada**, London, Ontario, Canada (Web: caa-aca.ca)

19–21

36th Spanish Congress on Acoustics Joint with 2005 Iberian Meeting on Acoustics, Terrassa, Barcelona, Spain (Sociedad Española de Acústica, Serrano 114, 28006 Madrid, Spain; Fax: +34 914 117 651; Web: www.ia.csic.es/sea/index.html)

25–26

***Autumn Conference 2005 of the UK Institute of Acoustics**, Oxford, UK (Web: www.ioa.org.uk)

November 2005

4–5

***Reproduced Sound 21**, Oxford, UK (Web: www.ioa.org.uk)

June 2006

05–06

6th European Conference on Noise Control (EURONOISE2006), Tampere, Finland (Fax: +358 9 7206 4711; Web: www.acoustics.hut.fi/asf)

26–28

9th Western Pacific Acoustics Conference (WESPAC 9), Seoul, Korea (Web: www.wespac9.org)

July 2006

3–7

***13th International Congress on Sound and Vibration (ICSV13)**, Vienna, Austria (Web: info.tuwienac.at/icsv13)

September 2006

13–15

***Autumn Meeting of the Acoustical Society of Japan**, Kanazawa, Japan (Acoustical Society of Japan, Nakaura 5th-Bldg., 2-18-20 Sotokanda, Chiyoda-ku, Tokyo 101-0021, Japan; Fax: +81 3 5256 1022; Web: www.asj.gr.jp/index-en.html)

July 2007

9–12

14th International Congress on Sound and Vibration (ICSV14), Cairns, Australia (e-mail: n.kessissoglou@unsw.edu.au)

August 2007

27–31

***Interspeech 2007** Antwerp, Belgium (e-mail: conf@isca-speech.org)

September 2007

2–7

19th International Congress on Acoustics (ICA2007), Madrid, Spain (SEA, Serrano 144, 28006 Madrid, Spain; Web: www.ica2007madrid.org)

9–12

ICA Satellite Symposium on Musical Acoustics (ISMA2007), Barcelona, Spain (SEA, Serrano 144, 28006 Madrid, Spain; Web: www.ica2007madrid.org)

June 2008

23–27

Joint Meeting of European Acoustical Association (EAA), Acoustical Society of America (ASA), and

Acoustical Society of France (SFA), Paris, France
(e-mail: phillipe.blanc-benon@ec-lyon.fr)

July 2008
28-1

***9th International Congress on Noise as a Public Health Problem**, Mashantucket, Pequot Tribal Nation (ICBEN 9, P.O. Box 1609, Groton CT 06340-1609, USA: Web: www.icben.org)

Preliminary Announcements

May 2006
15-19

***IEEE International Conference on Acoustics, Speech, and Signal Processing**, Toulouse, France.

August 2010
TBA

***20th International Congress on Acoustics (ICA2010)**, Sydney, Australia (Web: www.acoustics.asn.au)

BOOK REVIEWS

P. L. Marston

Physics Department, Washington State University, Pullman, Washington 99164

These reviews of books and other forms of information express the opinions of the individual reviewers and are not necessarily endorsed by the Editorial Board of this Journal.

Editorial Policy: *If there is a negative review, the author of the book will be given a chance to respond to the review in this section of the Journal and the reviewer will be allowed to respond to the author's comments. [See "Book Reviews Editor's Note," J. Acoust. Soc. Am. 81, 1651 (May 1987).]*

Vibrations of Shells and Plates, Third Edition

Werner Soedel

Marcel Dekker, Inc., New York, NY, 2004.

688 pp. Price \$185.00 (hardcover). ISBN 0-8247-5629-0

This is the third revised edition of the above book. It contains 21 chapters and covers aspects of theoretical development of the equations, free and forced vibration problems, methods of analysis, and complicating effects. The book is well written and rather comprehensive. It contains major revisions and/or additions to the second edition published more than 10 years earlier.

Chapter 1 covers an interesting introduction on the history of the development of shell and plate vibration problems. The fundamental equations of deep, thin shells are presented in the second chapter. Love's shell equations are derived. Treatment of shells with nonuniform thickness is also presented. This revised edition includes a new analytical treatment of finding the radii of curvature. The equations are specialized into specific shell geometries in Chap. 3. In particular, equations are developed for circular conical shells, circular cylindrical shells, and spherical shells. Chapter 4 offers a short, but important, treatment of curved beams and special cases of plate and shell vibrations.

Free vibration of shells and plates is the subject of Chap. 5. Problems that treat vibrations of simple beams, inplane as well as transverse vibrations of rectangular plates, and vibrations of circular plates and cylindrical shells are presented. Additional material found in this revised edition includes treatment of inplane vibration, which is important when piezoelectric materials are used, and solutions to the free vibration problem using power series methods.

Chapter 6 offers some of the simplified shell equations. Among the simplifications presented are those of the membrane shell theories, where bending moments are ignored, and the equations of Donnel and Mushtari, which are applicable to shallow shells. These simplified equations typically present a significantly reduced set of equations to treat special problems. They could be of importance to practicing engineers. The treatment of shallow shell structures including doubly curved rectangular panels is expanded greatly in Chap. 6.

Chapters 7 and 21 offer presentation of approximate solution techniques. These include the methods of Rayleigh, Ritz, finite difference, and finite elements as well as others. Expressions of strain energy which are quite important for many approximate techniques are also presented. These expressions are written for various types of shells in this revised edition.

Forced vibrations of shells by modal expansion are treated in Chap. 8. In addition to treatment of steady-state harmonic response, the chapter offers treatment of impact and impulsive loads. Treatments of some interesting initial value problems are included in this revised edition. The treatment of dynamic effects using Green's function is presented in Chap. 9.

Various complicating effects are presented in Chaps. 11–20. The effects of initial stresses on the shell vibrations are presented in Chap. 11. Treatment of shells having relatively high thickness is presented in Chap. 12. This treatment requires revising the fundamental assumption of Love leading to the derivation of new equations. Shear deformation and rotary inertia are to be included for such shells and/or plates. Chapter 13 addresses problems containing combinations of structures. These include structures with added mass as well as those made of more than one subsystem. Hys-

teresis damping is treated in Chap. 14, while composite structures are treated in Chap. 15. Treatment of sandwich structures is added to this chapter. Dealing with composite structures requires developing new equations for shells and/or plates. This was the subject of a new book by the reviewer.

Rotating structures are presented in Chap. 16. Of particular importance were rotating rings, rotating disks, and rotating circular cylindrical shells with the axis of rotation being the same as the axis of revolution. Some of these are additions found in the revised edition.

Thermal effects are discussed in Chap. 17. Treatment of shells on elastic foundation is presented in Chap. 18. This chapter is extensively revised in the third edition. Problems related to tire mechanics are addressed. In addition, problems addressing shells subjected to base excitation are also found in this revised edition. Interaction with gases and fluids is presented in Chap. 20. This addresses problems of submerged shell structures as well as fluid-filled ones.

Chapter 19 is an interesting discussion on the similitude and derivation of exact relations between scaled structures. This is of importance when tests are to be made on smaller scale structures for economic reasons. It is also important in building the rationale for nondimensionalization.

The book is an important reference to both researchers and practicing engineers who deal with the issue of vibration of shells and plates in their work. It is an excellent addition of knowledge to the engineering community and an important collection of specialized problems. This revised edition offers significant new material. One minor shortcoming of the book that should be stated is the drawings. They could have been improved in this revised edition to the high level of content reflected by the analytical treatment in the book.

MOHAMAD S. QATU
Ford Motor Company
20800 Oakwood Boulevard
Dearborn, Michigan, 48124

Experimental Acoustics Inversion Methods for Exploration of the Shallow Water Environment

A. Caiti, J.-P. Hermand, S. M. Jesus, and M. B. Porter (Editors)

Kluwer Academic Publishers, Netherlands, 2000.
293 pp. Price: \$127.00 (hardcover), ISBN: 0792363051.

Shallow water acoustics, with the complexities caused by boundaries and variations in the water column, is a broad and fertile area for basic and applied research. The field has made exciting progress when theoretical work has met real data in shallow water environments. The greatest challenges may be inverse problems where the unknown and time-varying environments are estimated through acoustics signals as a means of remote sensing. Such inverse problems are, though complex, necessary because it is impossible to measure directly all relevant acoustic parameters in the vast spatial and temporal spaces of shallow waters. This title is truly welcome by anyone who has an interest in shallow water acoustics.

While shallow water acoustics inversion has been an area of active

research, the subject is so diverse that no single book has been published to give a coherent treatment of its various subjects. Alternatively, this title provides a compilation of several interrelated subjects in the area under a single hardcover. It is a result of a workshop attended by about 20 researchers in the field, each covering their individual efforts relevant to the book title. Readers will find that actual shallow water conditions provide the focus for this publication, because all the papers are based on real data from shallow water regions.

The contents can be divided loosely into four subjects. They deal, respectively, with the effects of internal tides, source localization, effects of bottom vegetation, and inversion of sediment parameters.

The three papers concerned with the investigation of internal tides are all based on the same data sets, dealing with different aspects of the problem. Rodriguez *et al.* observes that in the presence of internal wave packets, propagation signals can occasionally experience a gain, possibly due to resonating effects. Their modeling of a particular internal wave packet and sound traveling through a packet is consistent with their assumption. To experimentally investigate such propagation gain effects as well as loss effects over range and frequency, it seems necessary to measure snapshots of internal wave fields.

Stephenan *et al.* report that variations observed in acoustic signals can be associated with different environmental effects. Specifically, they show that the influence of the bottom on their sound propagation data is sufficiently different from that of the internal tides so that simple, separate signal processing methods can be applied to deal with both problems.

Porter *et al.* bridge the subjects of internal tides and source localization. A notable feature here is that broadband signals over an extended period of time are used to investigate source tracking. The authors show that while narrow-band signals would have demonstrated large variations in arrival structures, broadband signals show stable arrival structure even in strong range-dependent scenarios. Also of interest is that fairly simple modeling, without considering all the complicated environmental variations, is shown to yield good tracking results.

On the subject of source tracking and localization, the paper by Ianniello and Tattersall analyzes experimental performance of a sophisticated multiline horizontal array system. Using simultaneously towed arrays at different depths, the authors report that both source range and depth can be determined using range-independent modeling. To achieve their results, they relied on the availability of good environmental data. Clearly, the range of applicability and the performance of such tracking arrays merit further study.

Two papers in the book cover acoustic interactions with vegetation on the sea bottom. Bozzano discusses the development and testing of a high-frequency backscatter system operating in the mega-Hz range. It also provides ample background and references on the subject of sea grass detection using high-frequency sonar. An interesting paper by Hermand *et al.* uses a simple experimental design to study sound propagation over sea grass. The goal of the investigation is to use sound as a remote sensing tool to monitor oxygen synthesis by sea grass. Contemporaneous measurements of temperature, sunlight, and CTD, as well as sound propagation, enable the authors to relate the abrupt and marked change in sound attenuation to the onset of photosynthesis in the sea grass. Furthermore, the authors are able to conclude that bubble layers formed from the photosynthesis are responsible for the sound attenuation and dispersion. Their work is an encouraging step forward to realizing the goal of using sound to monitor bubble concentration on the sea grass, hence providing a method to study photosynthesis of bottom vegetation.

The greatest number of papers in this collection concerns the inversion of sea bottom properties and parameters. Among all the uncertainties of the shallow water environment, the bottom is the more complex compared to the sea surface and the water column because of its diverse composition and the various processes controlling its variability. Therefore, it is not surprising that controversy continues concerning the appropriate governing equation used to describe the bottom as an acoustic medium. The theoretical paper by Buckingham argues for a viscous-fluid wave equation based on measured data showing that the sediment attenuation coefficient expressed in decibels has a linear relation to frequency. More details of the theory can be found in papers in the *Journal of the Acoustical Society of America*. Chotiros, however, argues for a poro-elastic model of sandy sediments. He

discusses the sensitivity of reflection loss as a function of grazing angle and frequency to various Biot parameters within the poro-elastic model and concludes that there exists the possibility to invert for some of the Biot parameters, such as porosity and grain density, using measurements of reflection loss. The appropriate wave equation to describe the sea bottom is an area of intense interest; research continues and progress has been made on these questions since the publication of this book.

Several papers report the use of sound propagation in shallow water environments to invert for bottom parameters. The experimental techniques range from using a narrow band source combined with vertical line arrays, to broadband explosive sources, to drifting buoys. A common theme in all methods is to take advantage of the amplitude and phase information of the propagating field, combined with modeling of the forward field based on different approximations to the environment, to estimate certain bottom parameters. Chapman *et al.* use imploding light bulbs as sound sources and vertical line arrays to measure broadband propagation. Their inversion is based on matched field processing by applying ray tracing to short range data. It is encouraging to read that the researchers were able to use a relatively straightforward method to infer bottom properties.

Abawi *et al.* give a systematic comparison of several matched field processing methods based on a set of shallow water propagation data collected by a tilted line array from a towed narrow-band source. This investigation is timely and important in order to transition matched field research into applications.

Rogers *et al.* summarize many years of experimental study of the Yellow Sea. There have been several experiments in the same general area over a span of many years, representing a persistent effort to understand bottom properties in one particular shallow water environment. Of particular interest is the reported result on the nonlinear dependence of bottom attenuation coefficient expressed in decibels versus frequency. Clearly, in the frequency band below 1000 Hz, more investigation is needed on the frequency dependence of sediment sound speed and attenuation.

Hermand *et al.* report research on geoacoustic inversion with drifting buoys. Several types of drifting buoys are tested over different sediments. A controlled source is used to transmit to these buoys. The abilities of these buoys with large dynamic ranges are articulated. This reviewer wishes there were more space devoted to the details of their inversion methods and a discussion of the success or failure of the inversion.

The paper by Caiti and Bergem is an ambitious effort to invert for a large number of bottom parameters from normal incidence backscatter data using a parametric source. The inverted parameters include not only the mean properties of the bottom such as mean sound speed and density, but also parameters controlling interface roughness and subbottom heterogeneity. Normal incidence data are attractive because they can be obtained over long tracks. If reliable inversions can be obtained using normal incident data, towed systems, such as the reported parametric sonar or chirp sonar, will be able to provide a practical method to collect bottom parameters over long tracks or large areas. However, this paper can only be considered an initial effort toward that goal because it does not offer a convincing case for the validity of the inversion results. To demonstrate the validity of the inversion, it is important to (1) establish the soundness of the forward model for the environment concerned, (2) independently measure with adequate resolution environmental parameters, and (3) assess the accuracy and associated error analysis of the inversion technique. The authors have yet to achieve these goals. It should be emphasized that their work is in an important area where increased effort ought to be placed.

While this book is not an introductory text for students, it is a useful reference for scientists and engineers working in the field of shallow water acoustics; through it an appreciation of several important areas of active research is gained. Readers should be aware that because of the limited number of papers, the subjects covered are by no means exhaustive for the area of acoustic inversion methods in shallow water environments.

DAJUN TANG
Applied Physics Laboratory
University of Washington
1013 NE 40th Street
Seattle, Washington 98105

BOOK REVIEWS

P. L. Marston

Physics Department, Washington State University, Pullman, Washington 99164

These reviews of books and other forms of information express the opinions of the individual reviewers and are not necessarily endorsed by the Editorial Board of this Journal.

Editorial Policy: *If there is a negative review, the author of the book will be given a chance to respond to the review in this section of the Journal and the reviewer will be allowed to respond to the author's comments. [See "Book Reviews Editor's Note," J. Acoust. Soc. Am. 81, 1651 (May 1987).]*

Experimental Acoustics Inversion Methods for Exploration of the Shallow Water Environment

A. Caiti, J.-P. Hermand, S. M. Jesus, and M. B. Porter (Editors)

*Kluwer Academic Publishers, Netherlands, 2000.
293 pp. Price: \$127.00 (hardcover), ISBN: 0792363051.*

Shallow water acoustics, with the complexities caused by boundaries and variations in the water column, is a broad and fertile area for basic and applied research. The field has made exciting progress when theoretical work has met real data in shallow water environments. The greatest challenges may be inverse problems where the unknown and time-varying environments are estimated through acoustics signals as a means of remote sensing. Such inverse problems are, though complex, necessary because it is impossible to measure directly all relevant acoustic parameters in the vast spatial and temporal spaces of shallow waters. This title is truly welcome by anyone who has an interest in shallow water acoustics.

While shallow water acoustics inversion has been an area of active research, the subject is so diverse that no single book has been published to give a coherent treatment of its various subjects. Alternatively, this title provides a compilation of several interrelated subjects in the area under a single hardcover. It is a result of a workshop attended by about 20 researchers in the field, each covering their individual efforts relevant to the book title. Readers will find that actual shallow water conditions provide the focus for this publication, because all the papers are based on real data from shallow water regions.

The contents can be divided loosely into four subjects. They deal, respectively, with the effects of internal tides, source localization, effects of bottom vegetation, and inversion of sediment parameters.

The three papers concerned with the investigation of internal tides are all based on the same data sets, dealing with different aspects of the problem. Rodriguez *et al.* observes that in the presence of internal wave packets, propagation signals can occasionally experience a gain, possibly due to resonating effects. Their modeling of a particular internal wave packet and sound traveling through a packet is consistent with their assumption. To experimentally investigate such propagation gain effects as well as loss effects over range and frequency, it seems necessary to measure snapshots of internal wave fields.

Stephanan *et al.* report that variations observed in acoustic signals can be associated with different environmental effects. Specifically, they show that the influence of the bottom on their sound propagation data is sufficiently different from that of the internal tides so that simple, separate signal processing methods can be applied to deal with both problems.

Porter *et al.* bridge the subjects of internal tides and source localization. A notable feature here is that broadband signals over an extended period of time are used to investigate source tracking. The authors show that while narrow-band signals would have demonstrated large variations in arrival structures, broadband signals show stable arrival structure even in strong range-dependent scenarios. Also of interest is that fairly simple modeling, without considering all the complicated environmental variations, is shown to yield good tracking results.

On the subject of source tracking and localization, the paper by Ianniello and Tattersall analyzes experimental performance of a sophisticated multiline horizontal array system. Using simultaneously towed arrays at

different depths, the authors report that both source range and depth can be determined using range-independent modeling. To achieve their results, they relied on the availability of good environmental data. Clearly, the range of applicability and the performance of such tracking arrays merit further study.

Two papers in the book cover acoustic interactions with vegetation on the sea bottom. Bozzano discusses the development and testing of a high-frequency backscatter system operating in the mega-Hz range. It also provides ample background and references on the subject of sea grass detection using high-frequency sonar. An interesting paper by Hermand *et al.* uses a simple experimental design to study sound propagation over sea grass. The goal of the investigation is to use sound as a remote sensing tool to monitor oxygen synthesis by sea grass. Contemporaneous measurements of temperature, sunlight, and CTD, as well as sound propagation, enable the authors to relate the abrupt and marked change in sound attenuation to the onset of photosynthesis in the sea grass. Furthermore, the authors are able to conclude that bubble layers formed from the photosynthesis are responsible for the sound attenuation and dispersion. Their work is an encouraging step forward to realizing the goal of using sound to monitor bubble concentration on the sea grass, hence providing a method to study photosynthesis of bottom vegetation.

The greatest number of papers in this collection concerns the inversion of sea bottom properties and parameters. Among all the uncertainties of the shallow water environment, the bottom is the more complex compared to the sea surface and the water column because of its diverse composition and the various processes controlling its variability. Therefore, it is not surprising that controversy continues concerning the appropriate governing equation used to describe the bottom as an acoustic medium. The theoretical paper by Buckingham argues for a viscous-fluid wave equation based on measured data showing that the sediment attenuation coefficient expressed in decibels has a linear relation to frequency. More details of the theory can be found in papers in the *Journal of the Acoustical Society of America*. Chotiros, however, argues for a poro-elastic model of sandy sediments. He discusses the sensitivity of reflection loss as a function of grazing angle and frequency to various Biot parameters within the poro-elastic model and concludes that there exists the possibility to invert for some of the Biot parameters, such as porosity and grain density, using measurements of reflection loss. The appropriate wave equation to describe the sea bottom is an area of intense interest; research continues and progress has been made on these questions since the publication of this book.

Several papers report the use of sound propagation in shallow water environments to invert for bottom parameters. The experimental techniques range from using a narrow band source combined with vertical line arrays, to broadband explosive sources, to drifting buoys. A common theme in all methods is to take advantage of the amplitude and phase information of the propagating field, combined with modeling of the forward field based on different approximations to the environment, to estimate certain bottom parameters. Chapman *et al.* use imploding light bulbs as sound sources and vertical line arrays to measure broadband propagation. Their inversion is based on matched field processing by applying ray tracing to short range data. It is encouraging to read that the researchers were able to use a relatively straightforward method to infer bottom properties.

Abawi *et al.* give a systematic comparison of several matched field processing methods based on a set of shallow water propagation data collected by a tilted line array from a towed narrow-band source. This investi-

gation is timely and important in order to transition matched field research into applications.

Rogers *et al.* summarize many years of experimental study of the Yellow Sea. There have been several experiments in the same general area over a span of many years, representing a persistent effort to understand bottom properties in one particular shallow water environment. Of particular interest is the reported result on the nonlinear dependence of bottom attenuation coefficient expressed in decibels versus frequency. Clearly, in the frequency band below 1000 Hz, more investigation is needed on the frequency dependence of sediment sound speed and attenuation.

Hernand *et al.* report research on geoacoustic inversion with drifting buoys. Several types of drifting buoys are tested over different sediments. A controlled source is used to transmit to these buoys. The abilities of these buoys with large dynamic ranges are articulated. This reviewer wishes there were more space devoted to the details of their inversion methods and a discussion of the success or failure of the inversion.

The paper by Caiti and Bergem is an ambitious effort to invert for a large number of bottom parameters from normal incidence backscatter data using a parametric source. The inverted parameters include not only the mean properties of the bottom such as mean sound speed and density, but also parameters controlling interface roughness and subbottom heterogeneity. Normal incidence data are attractive because they can be obtained over long tracks. If reliable inversions can be obtained using normal incident data, towed systems, such as the reported parametric sonar or chirp sonar,

will be able to provide a practical method to collect bottom parameters over long tracks or large areas. However, this paper can only be considered an initial effort toward that goal because it does not offer a convincing case for the validity of the inversion results. To demonstrate the validity of the inversion, it is important to (1) establish the soundness of the forward model for the environment concerned, (2) independently measure with adequate resolution environmental parameters, and (3) assess the accuracy and associated error analysis of the inversion technique. The authors have yet to achieve these goals. It should be emphasized that their work is in an important area where increased effort ought to be placed.

While this book is not an introductory text for students, it is a useful reference for scientists and engineers working in the field of shallow water acoustics; through it an appreciation of several important areas of active research is gained. Readers should be aware that because of the limited number of papers, the subjects covered are by no means exhaustive for the area of acoustic inversion methods in shallow water environments.

DAJUN TANG
Applied Physics Laboratory
University of Washington
1013 NE 40th Street
Seattle, Washington 98105

OBITUARIES

Søren Buus • 1951–2004



Søren Buus, a Fellow of the Society and a prominent researcher in speech and hearing science, died on 29 April 2004, after a nine-month illness. At the time of his death, he was Professor in the Department of Electrical and Computer Engineering and Director of the Communications and Digital Signal Processing Center at Northeastern University in Boston.

Professor Buus was born in Denmark on 29 January 1951 and developed an early interest in music and electronics. This led to an M.S. degree in 1976 from the Technical University of Denmark in Electrical Engineering and Acoustics. Soon thereafter, he came to Northeastern University in the United States, where he received his Ph.D. with Bertram Scharf in 1980 in experimental psychology. Early professional appointments were as a Senior Scientist at Northeastern, as a researcher designing Fender musical instruments at CBS, and as a Research Associate at Harvard with David Green. In 1986, he accepted a faculty position at Northeastern University, where he subsequently established the Hearing Research Laboratory of that University, and commenced a long-term collaboration with Mary Florentine, who directed the neighboring Communication Research Laboratory. The two laboratories jointly fostered interdisciplinary research with participation including students and faculty from engineering, audiology, speech, linguistics, and psychology.

Buus' research interests were in hearing and in speech perception, including the application of digital signal processing and microcomputer hardware and software for hearing rehabilitation, and he collaborated extensively with many well-known researchers in these fields. Among the scientists who worked with Buus at Northeastern were Robert Carlyon, Andrew Oxenham, Andrzej Miskiewicz, Hannes Müsch, and Tilmann Zwicker. Other activities included stints as guest researcher with Georges Canévet and Bertram Scharf at the Laboratory of Mechanics and Acoustics, CNRS Marseille, with Zwicker and Klump at the Technical University of Munich, with Namba and Kuwano at Osaka University, and with Poulsen at the Technical University of Denmark. Buus was a Fellow of the Acoustical Society of America and served on the ASA's Technical Committee on Psychological and Physiological Acoustics. He was also on the Editorial Board of *Hearing Research*.

Over the years, Buus has supervised, collaborated in, and participated in a wide variety of significant research projects. One recent project concerned the relation between the ability to discriminate differences among simple and complex sounds and the ability to understand speech in difficult listening situations, as when the speech is degraded by noise, reverberation, and/or limited bandwidth. Another project focused on temporal, spectral,

and across-ear integration of loudness and on how the loudness of moderate-level sounds is affected by preceding higher-level sounds. The long-term goal of this research has been the development of a general, quantitative model of the perception of tones, noise, and speech in both normal and impaired hearing.

Buus published many articles during his life, a good number of which can be found in this journal. One major theme, initiated with the paper "Release from masking caused by envelope functions" (*JASA*, December 1985) was the theoretical explanation of comodulation masking release, which is the improvement in the detection of a tone buried in a fluctuating noise seen when the noise level is increased by adding noise bands with similar envelopes at other spectral locations. According to Buus' theory, detection improves because the additional bands serve as cues that allow the listener to vary the weight assigned to the signal band over time and to listen primarily when the short-term masker energy is low; in other words, the cues promote "listening in the valleys."

Subsequent research at several different laboratories supports the hypothesis that listening in the valleys can probably explain most instances of comodulation masking release. Another research thrust, initiated by the paper "An excitation-pattern model for intensity discrimination" (*JASA*, December 1981, coauthored with Florentine), combined the notion of excitation patterns with signal-detection theory to model intensity discrimination and to show that discrimination must be performed by integration of information across all stimulated frequency-selective channels. The model permitted Buus, together with Klump, Gleich, and Langemann (*JASA*, July 1995) to develop and test the first quantitative model of auditory perception in a nonhuman species (the starling), tying together a wide variety of physiological and behavioral data for that species. The integration of information across independent frequency bands embodied in both models has been used by other authors in the development of physiologically based models of perception.

One may also note the recent work with Müsch (his doctoral student) on the use of statistical decision theory to predict speech intelligibility, with a sequence of papers published in *JASA*, starting with June 2001.

At Northeastern, Buus was a popular and effective teacher, who loved to explain things and to help others to understand. An accomplished musician, he was a recognized authority on the technical and perceptual aspects of guitars and guitar amplifiers. Many members of the Society, for example, may remember the invited presentation he gave with Tilmann Zwicker in June 1998 at the ASA/ICA meeting, on distortion as an artistic tool for guitar playing (*JASA*, Vol. 103, p. 2797).

Buus leaves his mother, three siblings, and their families, his wife (Mary Florentine, Matthews Distinguished Professor of Audiology at Northeastern), and their daughter (Julia Buus Florentine, a student at Herlufsholm High School in Denmark).

B. SCHARF
M. FLORENTINE

REVIEWS OF ACOUSTICAL PATENTS

Lloyd Rice

11222 Flatiron Drive, Lafayette, Colorado 80026

The purpose of these acoustical patent reviews is to provide enough information for a Journal reader to decide whether to seek more information from the patent itself. Any opinions expressed here are those of reviewers as individuals and are not legal opinions. Printed copies of United States Patents may be ordered at \$3.00 each from the Commissioner of Patents and Trademarks, Washington, DC 20231. Patents are available via the Internet at <http://www.uspto.gov>.

Reviewers for this issue:

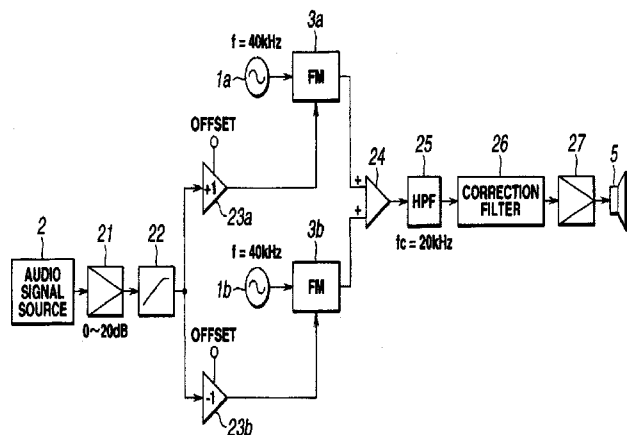
GEORGE L. AUGSPURGER, *Perception, Incorporated, Box 39536, Los Angeles, California 90039*
 JOHN M. EARGLE, *JME Consulting Corporation, 7034 Macapa Drive, Los Angeles, California 90068*
 SEAN A. FULOP, *California State University, Fresno, 5245 N. Backer Avenue M/S PB92, Fresno, California 93740-8001*
 IBRAHIM M. HALLAJ, *Wolf Greenfield & Sacks P.C., 600 Atlantic Avenue, Boston, Massachusetts 02210*
 JEROME A. HELFFRICH, *Southwest Research Institute, San Antonio, Texas 78228*
 HECTOR JAVKIN, *Department of Linguistics and Language Development, One Washington Square, San Jose, California 95192-0093*
 MARK KAHRIS, *Department of Electrical Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15261*
 NOLAND LEWIS, *ACO Pacific, Incorporated, 2604 Read Avenue, Belmont, California 94002*
 DAVID PREVES, *Starkey Laboratories, 6600 Washington Ave. S., Eden Prairie, Minnesota 55344*
 DANIEL R. RAICHEL, *2727 Moore Lane, Fort Collins, Colorado 80526*
 CARL J. ROSENBERG, *Acentech Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*
 NEIL A. SHAW, *Menlo Scientific Acoustics, Inc., Post Office Box 1610, Topanga, California 90290*
 KEVIN P. SHEPHERD, *Mail Stop 463, NASA Langley Research Center, Hampton, Virginia 23681*
 WILLIAM THOMPSON, JR., *Pennsylvania State University, University Park, Pennsylvania 16802*
 ERIC E. UNGAR, *Acentech, Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*
 ROBERT C. WAAG, *Univ. of Rochester, Department of Electrical and Computer Engineering, Rochester, New York 14627*

6,807,281

43.25.Lj LOUDSPEAKER AND METHOD OF DRIVING THE SAME AS WELL AS AUDIO SIGNAL TRANSMITTING/RECEIVING APPARATUS

Toru Sasaki *et al.*, assignors to Sony Corporation
 19 October 2004 (Class 381/111); filed in Japan 9 January 1998

During the past few years there has been considerable experimentation with loudspeakers that utilize an ultrasonic beam to create audible sound from empty air. None of this work is referenced in the patent at hand, but a thorough theoretical analysis of the disclosed scheme is included. Two



frequency-modulated signals are combined electronically or acoustically. The first is modulated by an audio signal and the second is modulated by an inverted version of the same audio signal. Corrective filtering is said to be greatly simplified, requiring no trigonometric functions. With a suitable array of piezoelectric transducers, a highly directional beam can be produced, with side lobes attenuated by more than 20 dB. The patent suggests that such

an arrangement can be used as a practical, low-cost source of "secret sound" for teleconferencing, museum exhibits, theme park rides, and private listening.—GLA

6,796,079

43.30.Nb SOLUBLE, MOVING, SOUND PRODUCING BAIT APPARATUS AND METHOD

J. Scott McCain, Minden, Louisiana
 28 September 2004 (Class 43/42); filed 15 November 2002

The patent author asserts that until his previous invention, "the lake, river, and ocean beds of the world were destined to become covered with an impervious layer of non-biodegradable plastic worm fishing lures." The author then argues that the use of bubbles will eliminate this waste. So, he proposes a mixture of citric acid and bicarbonate of soda, together with other attractants. When hydrated, CO₂ bubbles are generated, leading fish to the hook and the dinner table.—MK

6,804,167

43.30.Tg BI-DIRECTIONAL TEMPORAL CORRELATION SONAR

Anthony L. Scoca *et al.*, assignors to Lockheed Martin Corporation
 12 October 2004 (Class 367/89); filed 19 May 2003

A temporal correlation sonar is described that can continuously calculate the vector velocity of a surface ship. First and second pulses are radiated from the ship towards the ocean bottom. A planar array of hydrophones equally spaced around the perimeter of a square, i.e., a tetrad array, receives the echoes. Correlograms of the first pulse received at any hydrophone versus the second pulses received at the others are generated. Combining the known hydrophone separations and the times corresponding to the peak

amplitudes of these correlograms yields the velocity data. Using the data from a pair of hydrophones spaced along the ship's axis yields the forward velocity; similarly, athwart-ship velocity can be calculated. Since the velocity data stream generated is continuous, a simple time integration should produce ship-position information.—WT

6,813,220

43.30.Vh ACOUSTIC FENCE

Gary W. Hicks and Larry R. McDonald, assignors to Science Applications International Corporation
2 November 2004 (Class 367/136); filed 17 January 2003

An acoustic fence is realized by spacing a series of sonar modules, each with omni-directional radiation patterns, along some desired protective boundary. The modules may be spaced in relation to one another such that there is overlap in their coverage patterns. Each module collects range information on encroaching targets as a function of time. These data are relayed to a central processing unit via a telemetry link which then estimates the paths of these targets and makes threat level estimates relative to whatever the fence is supposed to be protecting.—WT

6,798,715

43.30.Wi BIOMIMETIC SONAR SYSTEM AND METHOD

John B. Harmon *et al.*, assignors to Neptune Technologies, Incorporated
28 September 2004 (Class 367/99); filed 9 July 2001

A portable sonar unit, capable of being carried by a diver or mounted on a small underwater vehicle, ensonifies submerged objects, digitizes the echo returns, and classifies and stores these digitized signals using an electronic processing procedure that is believed to emulate the process used by the bottlenose dolphin. Hence the name Biomimetic Sonar. The system then provides a probable identification of the object based on comparison with a library of similarly processed acoustic images.—WT

6,798,122

43.30.Yj LIGHTWEIGHT UNDERWATER ACOUSTIC PROJECTOR

Thomas R. Howarth *et al.*, assignors to The United States of America as represented by the Secretary of the Navy
28 September 2004 (Class 310/344); filed 5 November 2002

A planar grid array is realized by mounting a number of standard cylindrical "cymbal"-type flexensional piezoceramic driver elements within holes in a stiff mounting board. To mechanically isolate the drivers from the board, they are each surrounded at their outer rim by a soft rubber grommet. The elements can be electrically connected in various patterns for input impedance and directivity considerations.—WT

6,798,888

43.30.Yj MOUNT FOR UNDERWATER ACOUSTIC PROJECTOR

Thomas R. Howarth *et al.*, assignors to The United States of America as represented by the Secretary of the Navy
28 September 2004 (Class 381/162); filed 5 November 2002

In a variation of United States Patent 6,798,122, reviewed above, rather than each "cymbal"-type driver element having an individual isolation grommet and in turn being mounted in a common rigid support plate, here the driver elements are mounted in holes in a single rubber sheet which is then sandwiched between a pair of rigid acoustically transparent sheets of

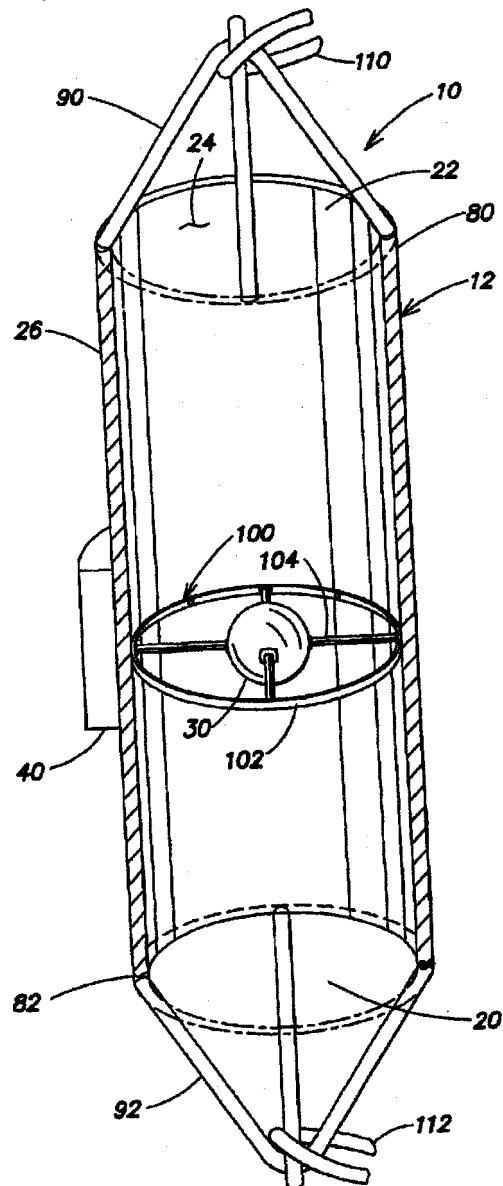
plastic such as Plexiglas™ with holes corresponding to the driver element positions.—WT

6,814,180

43.30.Yj MONOPOLE-DRIVEN UNDERWATER SOUND SOURCE

Hans Thomas Rosby and James H. Miller, assignors to The Board of Governors for Higher Education, State of Rhode Island and Providence Plantations
9 November 2004 (Class 181/120); filed 22 January 2002

An underwater sound source consists of a spherical piezoceramic monopole driver 30 supported at the center of a free-flooded metal or plastic tube 12 via an appropriate structure 100 and 104. The length of the tube, approximately 2 m, is chosen such that the source is resonant, in water, at



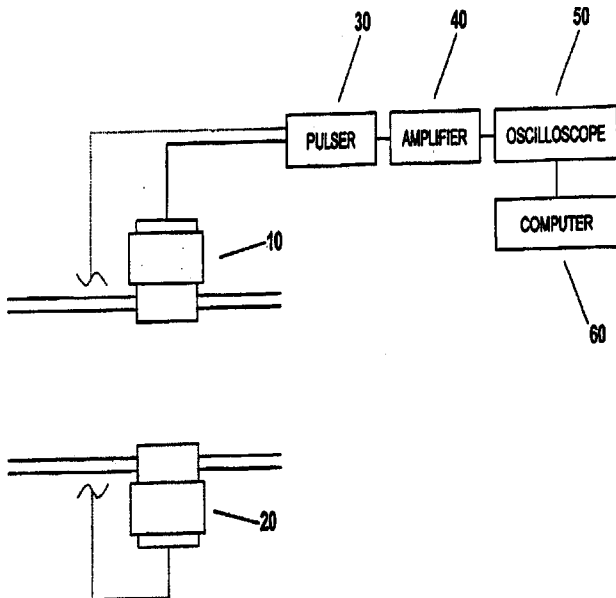
about 260 Hz. Harnesses 90 and 92 support mooring line 110 and weight line 112 respectively. Module 40, centered about the displacement node of the tube, houses the requisite electronics. Radiation into the surrounding medium is from the ends of the tube 20 and 22.—WT

6,796,195

43.35.Yb APPARATUS FOR DETERMINING PARTICLE SIZE

Malcolm J. W. Povey *et al.*, assignors to Baker Hughes, Incorporated
28 September 2004 (Class 73/865.5); filed 24 April 2003

This is an updated version of United States Patent 6,481,268 [reviewed in *J. Acoust. Soc. Am.* **113**(6), 2955 (2003)]. Measuring the size and concentration of small particles in fluids is an important consideration in many industrial processes, and several methods have been devised using ultrasonic imaging for this purpose. Some of these derive useful information from speckle, a characteristic that is usually regarded as a nuisance. The method described here introduces an ultrasonic signal into a fluid. A sensing transducer picks up the signal after it has been scattered by the particles and



feeds the resulting electrical signal to a computer. The computer is programmed to perform a Fourier transform, identify the speckle signal in the frequency domain, and relate certain characteristics of the speckle signal to particle size and distribution. The patent describes suitable algorithms in considerable detail and also includes a good summary of prior art, albeit with the proviso that "Applicants reserve the right to challenge the accuracy and pertinence of the cited references."—GLA

6,806,623

43.35.Yb TRANSMIT AND RECEIVE ISOLATION FOR ULTRASOUND SCANNING AND METHODS OF USE

David A. Petersen *et al.*, assignors to Siemens Medical Solutions USA, Incorporated
19 October 2004 (Class 310/334); filed 27 June 2002

Ultrasound imaging for echocardiography applications necessitates transducers with high volume-per-second rates for scanning. For real time imaging of moving structures, 20 or more two- or three-dimensional representations need to be generated each second. Large amounts of information are communicated from an ultrasound probe to an ultrasound base unit. The preferred embodiments described in this patent include methods and systems for isolating transmit and receive circuitry at an ultrasound transducer element. Separate electrodes or electrodes on opposite sides of a transducer element are connected to the separate transmit and receive paths or channels. Instead of high-voltage transmit and receive switching, the transducer element isolates the transmit channel from the receive channel. The transmit channel includes circuitry for limiting the voltage at one electrode during

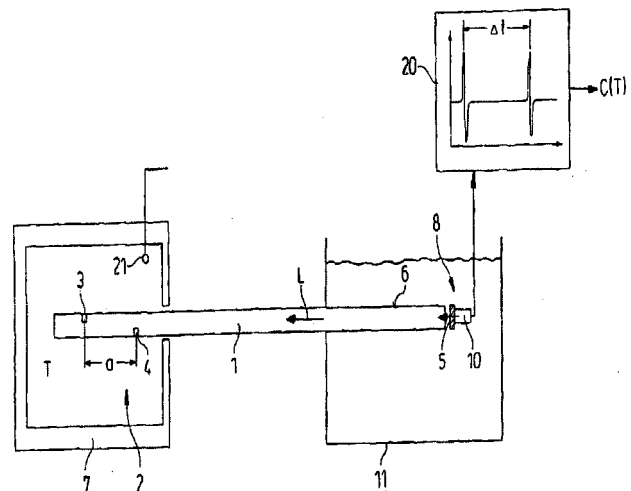
receive processing, such as a switch enable connecting the electrode to ground. The receive channel includes circuitry for limiting the voltage at an electrode during transmit processing, such as a diode clamp preventing voltage swings greater than diode voltage at the electrode. Limiting the voltage provides virtual grounding or a direct current for either the transmitting or receiving operation. A unipolar pulse may be generated on a transmit channel starting at one voltage and ending at a different voltage. For example, a pulse may start at zero voltage and end on a positive voltage, with a subsequent pulse starting at the positive voltage and ending at zero voltage. These mirrored unipolar transmit waveforms may be used for phase inversion imaging, such as adding responsive received signals for isolating harmonic information.—DRR

6,810,742

43.35.Zc METHOD OF AND AN APPARATUS FOR DETERMINING THE SPEED OF SOUND IN A MATERIAL

Martin Sauerland, assignor to SMS Meer GmbH
2 November 2004 (Class 73/597); filed in Germany 29 May 2002

The patent first explains how known techniques can monitor the wall thickness of pipe or tubing during production if the speed of sound through the material at the temperature of fabrication can be specified. Interpolating from standardized tables may not be accurate enough. The arrangement shown provides a simple, practical test technique. Sample 1 is a solid rod made from the material to be tested. Notches 3 and 4 are milled into the rod



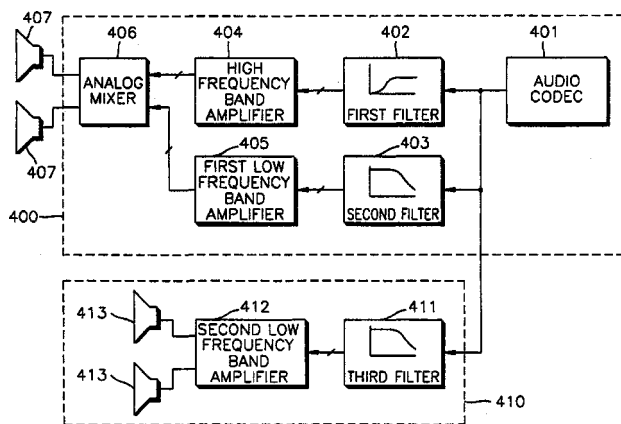
to act as reflectors. The notched end of the rod is placed inside furnace 7 and heated to the desired temperature, which may be greater than 600 °C. The opposite end of the rod is cooled by immersion in tank 11. Ultrasonic test head 10 generates a pulse and then receives reflections from the two notches. Since the distance between the notches is known, the time interval between the two reflections can be used to calculate sound speed.—GLA

6,813,528

43.38.Ar APPARATUS AND METHOD FOR OUTPUTTING AUDIO SIGNAL OF LAPTOP COMPUTER COUPLED WITH DOCKING STATION

Sung-dong Yang, assignor to Samsung Electronics Company, Limited
2 November 2004 (Class 700/94); filed in the Republic of Korea 9 March 1998

The patent deals with some of the challenges and opportunities of enhancing the audio performance of a laptop computer when it is coupled to a larger system via a so-called docking station. While retaining the small



stereo loudspeakers built into the basic laptop, the docking station supplies its own audio enhancements through the use of additional loudspeakers, primarily to extend the low-frequency response.—JME

6,812,617

43.38.Bs MEMS DEVICE HAVING A FLEXURE WITH INTEGRAL ELECTROSTATIC ACTUATOR

Thomas Wayne Ives, assignor to Hewlett-Packard Development Company, L.P.

2 November 2004 (Class 310/309); filed 18 April 2002

The patent describes a clever electrostatically actuated, micromechani-



cal "spine." The device looks neat, but no force or displacement capabilities are given. Probably just a sketchbook curiosity.—JAH

6,812,814

43.38.Bs MICROELECTROMECHANICAL (MEMS) SWITCHING APPARATUS

Qing Ma *et al.*, assignors to Intel Corporation

2 November 2004 (Class 333/262); filed 7 October 2003

This patent describes the layout of a tiny electrostatically actuated switch suitable for rf switching. Several embodiments and fabrication methods are described, with typical connections being made by a $5 \times 5 \times 20 \mu\text{m}$ gold bar, which should result in reasonable contact resistances. The focus of the inventors is on the geometry of the switch—no performance figures are given.—JAH

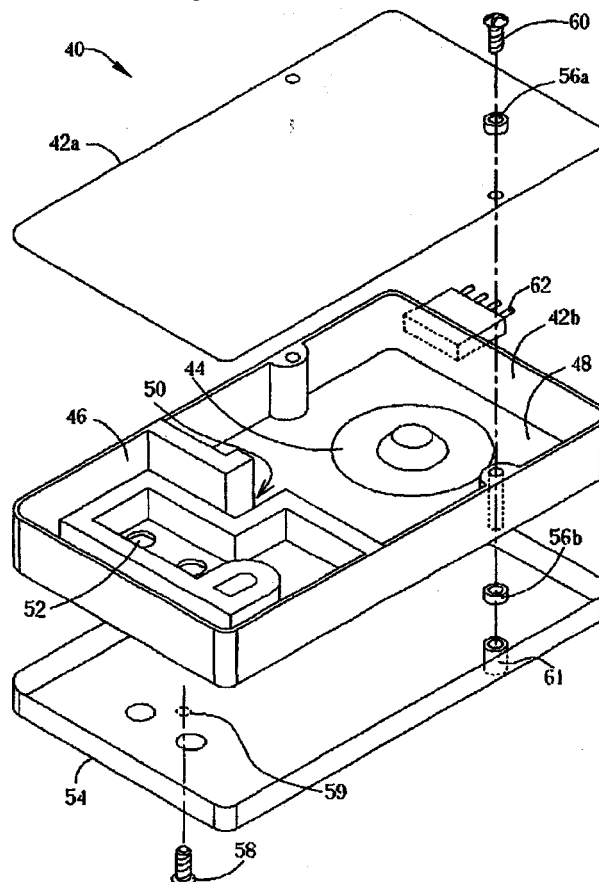
6,798,654

43.38.Dv WOOFER MODULE OF A PORTABLE COMPUTER

Hung-Yue Chang *et al.*, assignors to Wistron Corporation

28 September 2004 (Class 361/686); filed in Taiwan, Province of China 26 September 2002

The quality of sound produced by, even so-called, multimedia portable computers is in many cases a joke. Despite the inherent problems with trying to move lots of air to produce better quality sound from these devices (due to the minuscule space to place loudspeakers), the need by the manufacturer to use the cheapest transducers known to man in order to be cost



competitive, and the need for the user to suspend disbelief, many inventors still are toiling in the manner of Sisyphus. This patent describes a small ported enclosure that can attach to one of the computer's expansion ports or slots. One assumes that the stiffness of the enclosure is sufficient for the need and that the transducer is up to the task.—NAS

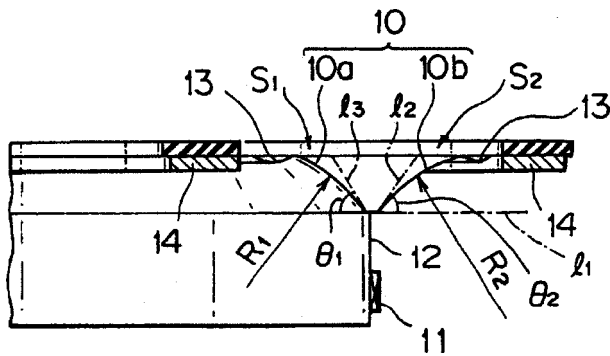
6,804,370

43.38.Dv SPEAKER CAPABLE TO PLAYBACK IN WIDE FREQUENCY RANGE

Hideaki Sugiura, assignor to Pioneer Corporation; Tohoku Pioneer Corporation

12 October 2004 (Class 381/424); filed in Japan 29 March 2002

Cone-type electrodynamic loudspeakers are known to have resonances. One of the notable resonances is that of the cone. The patent uses two diaphragms **10a** and **10b** that have different radii of curvature and that also have the benefit of different resonant frequencies. Somehow, by "equal-



izing areas" of the inner diaphragm, S1 for 10a, and of the outer diaphragm, S2 for 10b, the sound pressure levels are equalized with each other. The patent states that this compound cone, with the equalized area diaphragms, results in a smoother high-frequency response.—NAS

6,812,607

43.38.Dv AUXILIARY COUPLING DISK FOR MINIATURE VIBRATION VOICE-COIL MOTORS

Chien-Chung Sun, assignor to Risun Expanse Corporation
2 November 2004 (Class 310/81); filed 5 August 2003

This patent discloses a method for increasing the vibration output from an electric motor of limited size, through the use of a magnetic, eccentric rotor. The inventor's concept of small is "less than 20 mm diameter."—JAH

6,811,385

43.38.Fx ACOUSTIC MICRO-PUMP

Daniel R. Blakley, assignor to Hewlett-Packard Development Company, L.P.
2 November 2004 (Class 417/413.2); filed 31 October 2002

This patent describes a peristaltic fluid pump having an inner diameter of tens of microns created by a sequence of piezoelectric rings surrounding a tubular passage. There is a good deal of discussion of the preferred positions of the piezoelectric rings but next to nothing on the physics of what is going on in the channel. There is also no discussion of the amplitude dependence of what would apparently be a nonlinear effect, leading one to suspect that the device has not actually been studied.—JAH

6,812,618

43.38.Fx CONTROL APPARATUS FOR VIBRATION TYPE ACTUATOR

Tadashi Hayashi, assignor to Canon Kabushiki Kaisha
2 November 2004 (Class 310/316.01); filed in Japan
28 December 2000

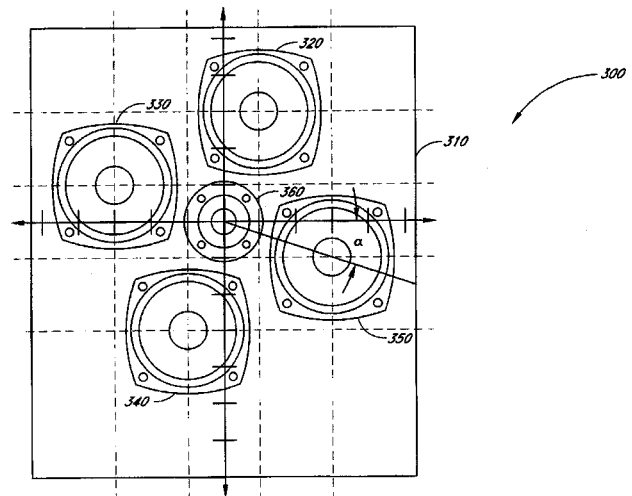
This patent describes several control circuits for piezoelectric traveling-wave-type actuators. As such it belongs in the field of electronics. No mention is made of the application of the circuits to MEMS actuators, though it most certainly could be done.—JAH

6,801,631

43.38.Hz SPEAKER SYSTEM WITH MULTIPLE TRANSDUCERS POSITIONED IN A PLANE FOR OPTIMUM ACOUSTIC RADIATION PATTERN

Donald J. North, Los Angeles, California
5 October 2004 (Class 381/336); filed 24 May 2000

Loudspeakers are usually used in rectangular rooms. It follows that first-order reflections from a loudspeaker to a listener must be influenced by the horizontal and vertical directional characteristics of the loudspeaker. Therefore, if a simple square array is rotated slightly as shown, then the



effective center-to-center spacing is reduced in the horizontal and vertical planes, attenuating off-axis lobes up to a higher Nyquist frequency. In theory, it all makes sense.—GLA

6,809,465

43.38.Hz ARTICLE COMPRISING MEMS-BASED TWO-DIMENSIONAL E-BEAM SOURCES AND METHOD FOR MAKING THE SAME

Sungho Jin, assignor to Samsung Electronics Company, Limited
26 October 2004 (Class 313/310); filed 24 January 2003

This is a description of a MEMS device that allows you to make and steer a large number of electron beams for use in a plasma panel or an x-ray source. The approach is to make an array of tips that can be steered electrically. There is no mention of how the device would be constructed or whether the various materials are process compatible.—JAH

6,792,120

43.38.Ja AUDIO SIGNAL ENHANCEMENT AND AMPLIFICATION SYSTEM

Jonathan M. Szenics, Flemington, New Jersey
14 September 2004 (Class 381/120); filed 23 February 2000

Motional feedback has been known for almost 50 years. According to the inventor, "tube amplifiers are still the gold standard by which any new amplification/signal processing device is judged." His solution is to epoxy a piezoelectric sensor to a speaker cone. That signal is then amplified with a

coil rubs against one of the pole pieces. Moreover, the inner flange of the suspension can be conveniently tuned to supply additional energy at very high frequencies.—GLA

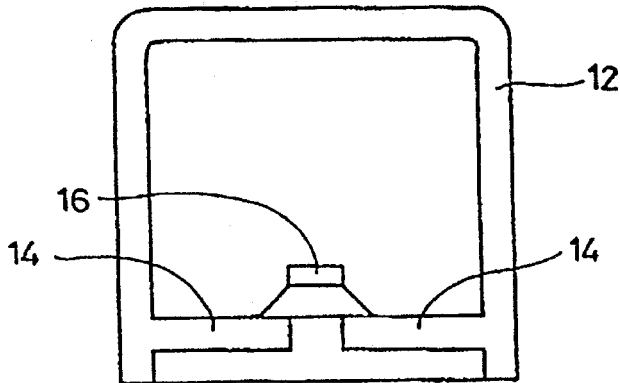
6,807,284

43.38.Ja SPEAKER AND SPEAKER DEVICE

Takeshi Nakamura, assignor to Murata Manufacturing Company, Limited
19 October 2004 (Class 381/423); filed in Japan 28 September 1998

The patent describes a device that is said to emit nondirectional, low-frequency waves. This is done by having a loudspeaker 16 mounted in a case 12 such that the volume of air enclosed is airtight. "By storing the speaker unit inside of a very small case, the resonant frequency of the speaker is significantly increased, enabling the fidelity of the speaker to be improved by generating sound at frequencies below the resonant frequency.

10



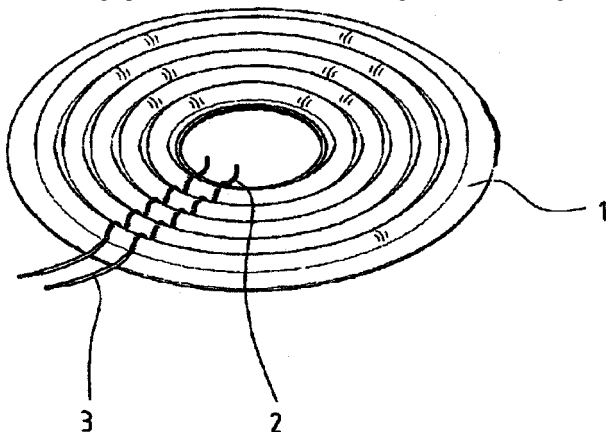
By providing the speaker unit on a floor side of the case, it is possible to cut off sound emitted by the speaker unit to the outside thereof, so that only sound generated by the expansion and contraction of the case is emitted." Which, in the least, doesn't sound very efficient.—NAS

6,810,988

43.38.Ja SPEAKER'S DAMPER WITH LEAD WIRE AND GUIDE SLEEVE

Yen-Chen Chan, Taipei Hsien, Taiwan, Province of China
2 November 2004 (Class 181/171); filed 2 December 2002

Centering spider 1 is a conventional corrugated, woven design. Lead



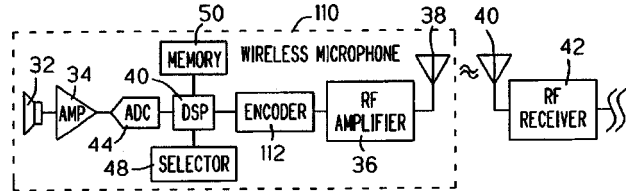
wires are threaded through little loops, like basting, to keep them separated and restrain unwanted vibrations.—GLA

6,810,125

43.38.Kb MICROPHONE EMULATION

Doran Oster et al., assignors to Sabine, Incorporated
26 October 2004 (Class 381/111); filed 4 February 2002

For musicians who are into digital production techniques the notion of a microphone simulator is not new. Just as the impulse response of classic equalizers and other effects generators of past years can be measured and



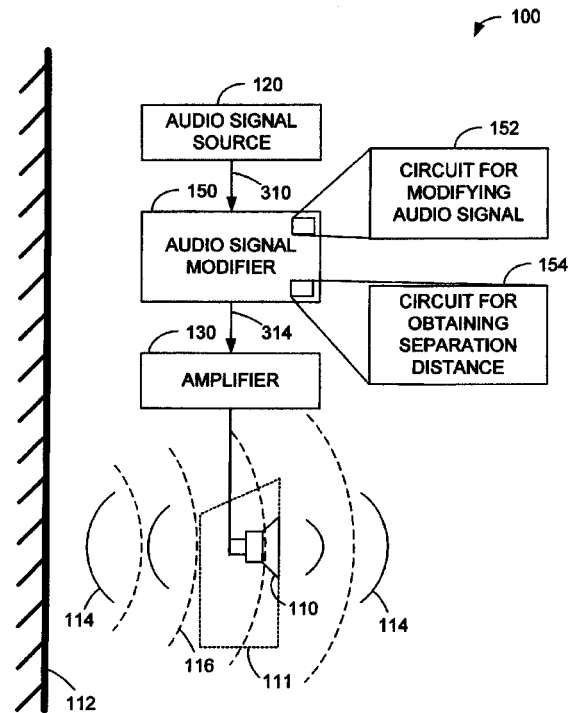
used to simulate the original devices, so likewise can the response of classic microphones be duplicated for similar purposes. There is absolutely nothing new in this basic premise, but one of the claims of the patent describes its application in the context of a standard wireless microphone channel. Why? Perhaps to establish some kind of primacy in the field of live music entertainment, in which wireless microphones are widely used by vocalists on the move.—JME

6,801,628

43.38.Lc SYSTEM AND METHOD FOR ADJUSTING FREQUENCY RESPONSE CHARACTERISTICS OF A SPEAKER BASED UPON PLACEMENT NEAR A WALL OR OTHER ACOUSTICALLY-REFLECTIVE SURFACE

James Thiel, assignor to Thiel Audio Products
5 October 2004 (Class 381/56); filed 14 May 2001

There are several, if not more, systems commercially available today that take into account what is called the "room acoustics." This patent describes a device 150 that is inserted into the signal chain before a power amplifier 130 to adjust the signal 310 from signal source device 120. The distance from the wall is the main input parameter to 152, which is obtained



via manual entry or by automatic means from 154. The object is to reduce the peak in response that one can find at the low end due to the wall reflection. The equations used vary as a function of distance, with the coefficients changing at a distance of approximately 2.7 m. An example of an implementation, with design equations, is provided.—NAS

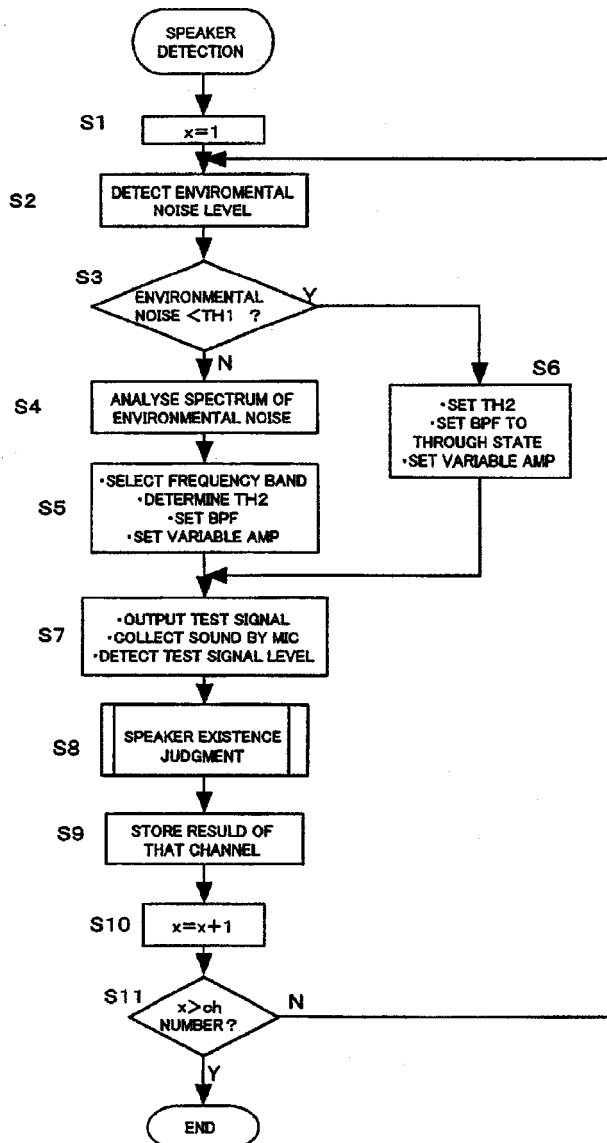
6,813,577

43.38.Lc SPEAKER DETECTING DEVICE

Hajime Yoshino and Kazuya Tsukada, assignors to Pioneer Corporation

2 November 2004 (Class 702/111); filed in Japan 27 April 2001

Many surround sound systems include a test mode in which pink noise is sequentially fed to all six loudspeaker channels. The flow chart shows how a computer program can augment this simple procedure to automati-



cally check whether each speaker is operational—not for relative level, polarity, or delay—just whether it is alive or dead. If there is any memory left over, another program might be added to let the user know if the equipment is actually turned on when the pilot light is lit.—GLA

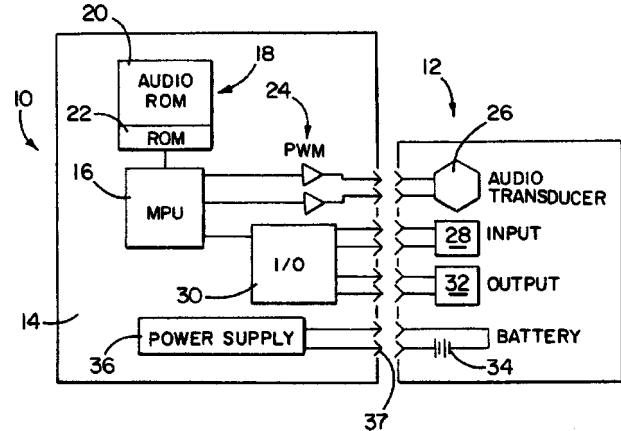
6,801,815

43.38.Md SOUND AND IMAGE PRODUCING SYSTEM

Andrew S. Filo and David G. Capper, assignors to Hasbro, Incorporated

5 October 2004 (Class 700/94); filed 28 July 2000

In the cut-throat toy business, the recording and playback of sounds for toys and similar entertainment devices is often the one important component in the device that generates the sale of same. The present invention appears to be an outgrowth and enhancement of the assignee's Hit Clips product line, albeit one that could have better sound quality (this information comes from the reviewer's resident expert on such devices). Player 12, which, in the several embodiments described appears to be compact in size, appears to



be a "dumb" device in that it has little or no personality or audio functionality until cartridge 10 is plugged in. In other embodiments, device 10 can be connected to other equipment to download new sounds and personality. The patent describes in minute detail the construction of the cartridges and some of the forms that the players can take. Further, it states that "users may purchase and collect cartridges; users may also trade cartridges amongst themselves; high school students may clip players to a zipper on their backpacks." The expert consulted is not quite that old.—NAS

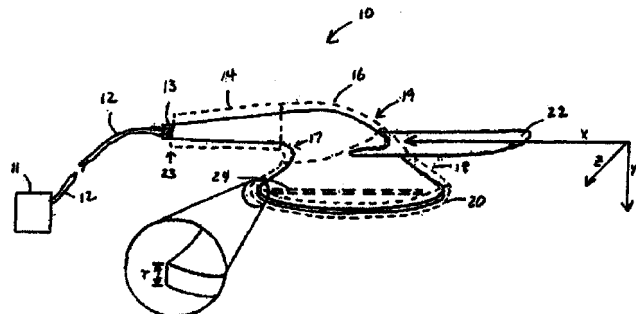
6,810,987

43.38.Si EARBUD HEADSET

Frederick P. DeKalb, assignor to Plantronics, Incorporated

2 November 2004 (Class 181/129); filed 6 December 2002

The earbud shown is intended to provide enhanced sound quality and positional stability in ears of various sizes. Novel features include a thin



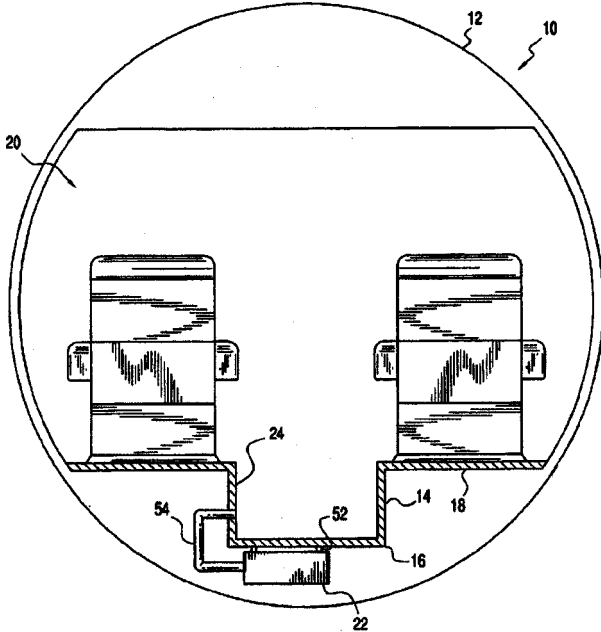
edge portion 20, which is said to allow acoustic coupling to deeper areas of the user's ear, and positioning member 22, which contacts an upper concha of the ear.—GLA

6,801,633

43.38.Tj SUBWOOFER ASSEMBLY

Dennis A. Tracy, Culver City, California
 5 October 2004 (Class 381/389); filed 25 January 2001

Any passenger on a commercial aircraft who has listened to the audio system via headphones realizes that the low-frequency reproduction is even poorer than that of the mid frequencies (the high frequencies being virtually nonexistent due to the ambient noise in the cabin). In fact, high-quality sound to accompany the in-flight entertainment is not only not available, today it is an oxymoron. "A need, therefore, exists for a speaker assembly providing high-quality sound, while also meeting the size and weight requirements of an aircraft." This patent sets out to provide some improve-



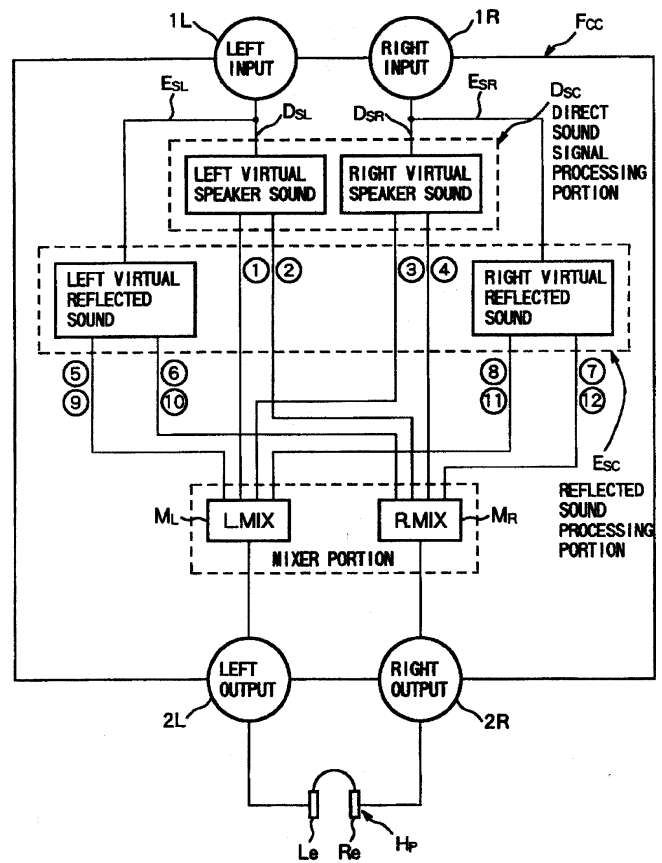
ment at subwoofer frequencies. By using a subwoofer enclosure 22 mounted below the aisle 16 in the aircraft fuselage 12, the low frequencies produced by this device within the passenger compartment are said to be improved. One supposes the improvement here is that of the sound level. Great, so you will soon not only have the pleasure of listening to the multiple cell phone conversations around you, but they will be accompanied by the subwoofer component of the video that you are not watching.—NAS

6,801,627

43.38.Vk METHOD FOR LOCALIZATION OF AN ACOUSTIC IMAGE OUT OF MAN'S HEAD IN HEARING A REPRODUCED SOUND VIA A HEADPHONE

Wataru Kobayashi, assignor to OpenHeart, Limited; Research Network, a Limited Responsibility Company
 5 October 2004 (Class 381/17); filed in Japan 30 September 1998

This patent describes a method for making headphones sound like loudspeakers by simulating both direct sound and first-order reflections. The



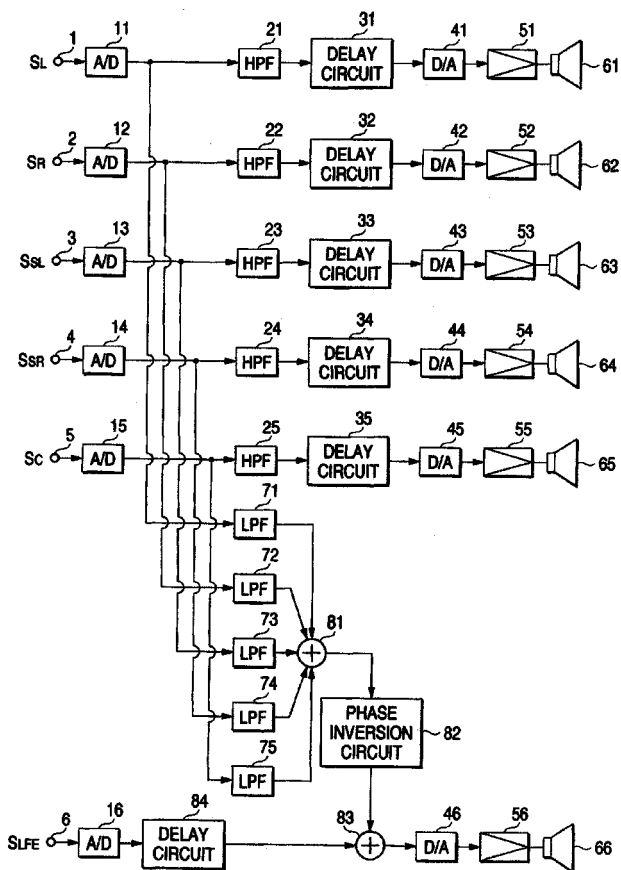
inventors maintain that an audio signal above a certain critical frequency can be positioned at any virtual location through a combination of delay and comb-filtering.—GLA

6,804,361

43.38.Vk SOUND SIGNAL PLAYBACK MACHINE AND METHOD THEREOF

Shintaro Hosoi and Hiroyuki Hamada, assignors to Pioneer Corporation
 12 October 2004 (Class 381/99); filed in Japan 12 June 2001

Most conventional frequency-dividing networks produce substantial delay distortion in their summed response. A number of alternative constant-voltage circuits have been proposed, but are largely unused because of unwanted side effects. However, in 1985 Vanderkooy and Lipshitz disclosed a clever method of realizing linear phase response by introducing time offset into the high-frequency channel, and this approach is also used in the circuit shown. In this case, the inventors are concerned with the crossover between multiple surround sound speakers and a common subwoofer. To at least one



reader, the problem of synchronizing all five speakers with an off-center subwoofer is not dealt with convincingly.—GLA

6,799,465

43.38.Yn SHOCK WAVE VIBRATION GENERATOR FOR MEASURING OR TESTING OR IMAGING PURPOSES

Stephen Bruce Berman, Austin, Texas
5 October 2004 (Class 73/590); filed 16 July 2002

Those familiar with shock wave testing may find this patent informative. However, this reviewer finally gave up trying to make sense of sentences like, "In one embodiment of the present invention, a method utilizes one or more shock wave vibratory forces produced by one or more cavitating spaces within one or more fluid substances within one or more housings which is therefore an improvement of the art of methods for producing vibrator devices that are particularly adapted for producing and applying non-shock wave vibratory forces, for example, any testing purposes or any measuring purposes."—GLA

6,813,218

43.38.Zp BUOYANT DEVICE FOR BI-DIRECTIONAL ACOUSTO-OPTIC SIGNAL TRANSFER ACROSS THE AIR-WATER INTERFACE

Lynn T. Antonelli and Fletcher Blackmon, assignors to The United States of America as represented by the Secretary of the Navy
2 November 2004 (Class 367/3); filed 6 October 2003

A floating buoy consists of a hollow shell having a set of underwater acoustic transducers mounted against its lower portion while the upper portion features a dome-shaped retro-reflective coating that can be vibrated in accordance with received acoustic signals to retro-reflect impinging in-air

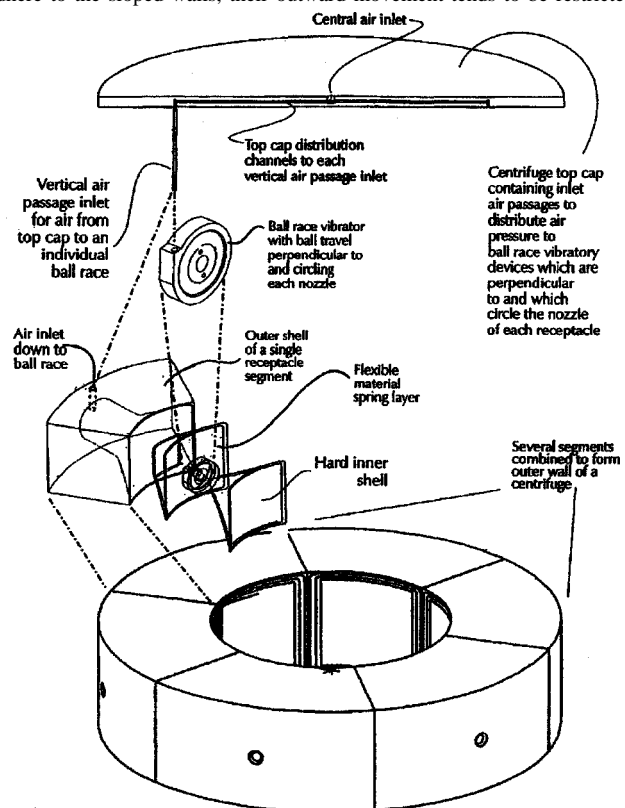
laser signals back to a remote receiver. Alternatively, the upper portion will vibrate in response to an incident in-air laser control signal. Consequently, the buoy functions as a system for either translating in-air optical signals to underwater acoustic signals or translating underwater acoustic signals to in-air optical signals.—WT

6,805,805

43.40.Ey SYSTEM AND METHOD FOR RECEPTACLE WALL VIBRATION IN A CENTRIFUGE

Curtis Kirker and Berkeley F. Fuller, assignors to Phase Incorporated
19 October 2004 (Class 210/781); filed 13 August 2002

In a centrifuge, the centrifugal force achieved by spinning drives the heavier materials outwards from the center of the centrifuge. The heavier materials contact the sloped and converging walls of collecting receptacles leading to exit outlets or nozzles. Because such heavier materials tend to adhere to the sloped walls, their outward movement tends to be restricted.



Vibrating the centrifuge's receptacle tends to loosen these materials from the walls and thus to promote their egress from the centrifuge. A means is provided here to vibrate the centrifuge receptacle through the use of a vibratory device communicating with the receptacle wall.—DRR

6,796,618

43.40.Kd METHOD FOR DESIGNING LOW VIBRATION OMNI-DIRECTIONAL WHEELS

Donald Barnett Harris, Arlington, Virginia
28 September 2004 (Class 301/501); filed 18 October 2002

Most omni-directional vehicles use wheels that feature rollers positioned about the periphery of the wheel, with the roller axes at some angle relative to the wheel's primary plane and axis. As the wheel turns, the area contacting the ground shifts across the surface of each roller and the portion of the roller that supports the vehicle's weight changes. With conventional

roller designs, the compliance that contacts the ground changes, resulting in changing vertical deflection of the vehicle and thus in vibration. The present patent proposes roller designs that avoid this problem by maintaining constant ground-contacting compliance.—EEU

6,789,025

43.40.Le CYCLIC TIME AVERAGING FOR MACHINE MONITORING

Johannes I. Boerhout, assignor to SKF Condition Monitoring, Incorporated
7 September 2004 (Class 702/56); filed 4 December 2002

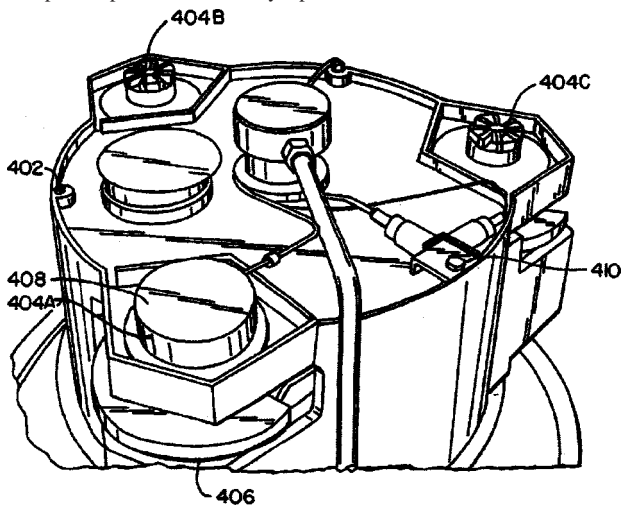
A common means of monitoring the condition of a machine uses vibration sensors attached to the machine, processing of the signals from the vibration transducer, and computer-based analysis of these signals to determine the contribution of bearing wear, shaft irregularities, etc., before the catastrophic failure of the machine. Here, a set of N vibration amplitude values is accumulated from which a set of n averaged amplitude values, separated by period M , is obtained, with $n < N$. A novel feature is that some of the n values do not correspond to any of the N values. A very novel part of the patent is that the final two figures referenced in the patent text do not appear in the patent. Figure 4 in this patent is identical to Figure 1 in United States Patent 6,792,360, seen with the review of that patent below.—NAS

6,789,422

43.40.Le METHOD AND SYSTEM FOR BALANCING A ROTATING MACHINERY OPERATING AT RESONANCE

Talmadge D. Ward, Jr., assignor to United States Enrichment Corporation
14 September 2004 (Class 73/462); filed 21 December 2001

This patent describes a method to speed the balancing of multishaft rotating equipment where the shafts are not necessarily aligned along a common axis of rotation. The patent presents a clear and useful tutorial on the various methods of shaft balancing, vibration analysis of machines, rotor dynamics, and the rotor influence coefficient test. An example of the method for a specific piece of machinery, specific transducers, and data collection



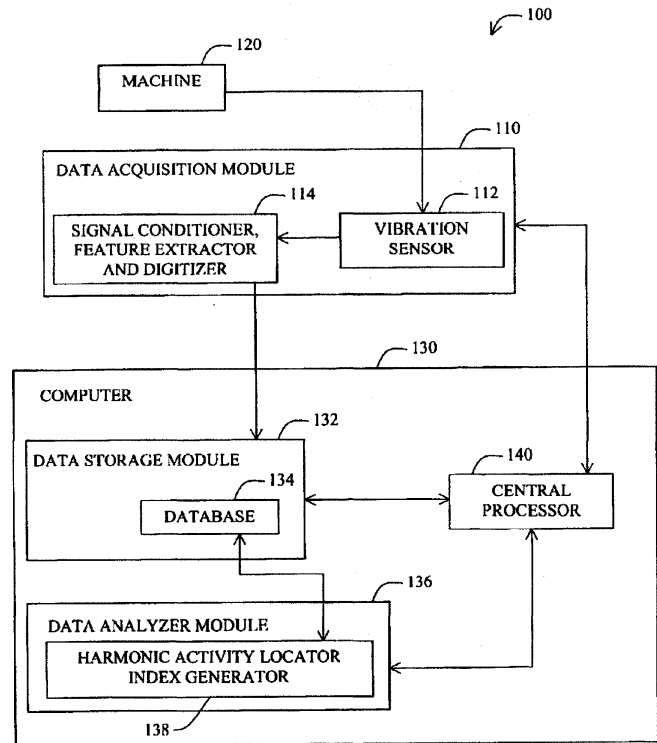
and analysis equipment is presented. The method uses a proximity transducer 408 on one shaft to provide rotational velocity data and a pair of vibration transducers 410 mounted perpendicular to one another. A detailed procedure is presented that purports to decrease the time to balance and, one would assume, the accuracy of the balancing.—NAS

6,792,360

43.40.Le HARMONIC ACTIVITY LOCATOR

Adrianus J. Smulders *et al.*, assignors to SKF Condition Monitoring, Incorporated
14 September 2004 (Class 702/35); filed 4 December 2002

This patent describes a computer-based method of processing vibration data in the frequency domain that uses a harmonic activity index based on frequency spectrum information and that of, one supposes, a known-defects harmonic series. It is written in classic "patenesque," stating in the claims that estimated values K and R are indicative of the total energy in the spectrum,



that a derived value I_{HAL} is based in part on K and R , and that using I_{HAL} and the fundamental frequency of the component defect it may be determined whether the vibrations of the machine are caused by the defect. There are 33 other claims, too.—NAS

6,801,864

43.40.Le SYSTEM AND METHOD FOR ANALYZING VIBRATION SIGNALS

Niel Miller, assignor to AB SKF
5 October 2004 (Class 702/56); filed 15 November 2001

Identification of defects or damage in rolling element bearings is accomplished by filtering vibration data in bands that are determined in part on the basis of the shaft's rotation speed. The bands are arranged to encompass signals that correspond to defects in the rotating inner ring, the rotating outer ring, and the rotating elements. User-defined alarm levels are determined by comparing the signals to a noise floor.—EEU

6,807,866

43.40.Le TRANSDUCER OF THE VIBRATION TYPE, SUCH AS AN ELECTROMECHANICAL TRANSDUCER OF THE CORIOLLIS TYPE

Wolfgang Drahm and Alfred Rieder, assignors to
Endress+Hauser Flowtec AG
26 October 2004 (Class 73/761.355); filed 22 February 2002

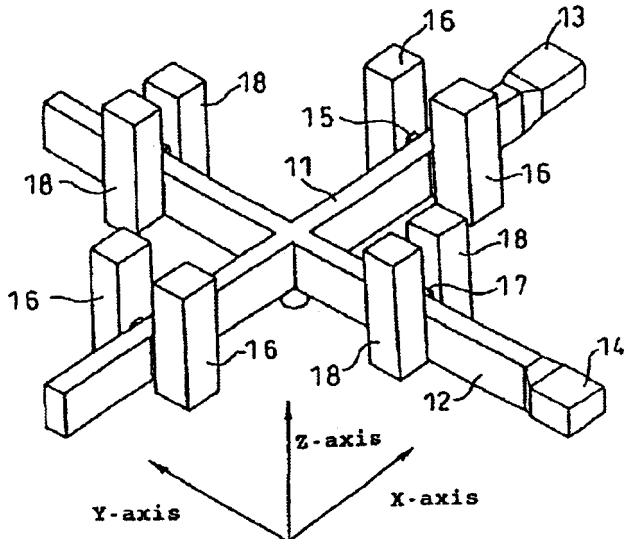
Coriolis mass flow meters are well known. Such a flow meter basically consists of a curved tube, which is made to vibrate at a resonance in its plane. Coriolis forces, which depend on the mass flow rate, then cause out-of-plane vibrations and these are sensed to obtain a measure of the flow rate. The present patent describes a particular configuration of the curved tube that results in minimal vibration transmission to its end supports.—EEU

6,811,630

43.40.Sk ULTRASONIC VIBRATION METHOD AND ULTRASONIC VIBRATION APPARATUS

Morio Tominaga and Shinji Iwahashi, assignors to Sony
Corporation
2 November 2004 (Class 156/73.1); filed in Japan
20 November 2000

This is a method for exciting a composite ultrasonic vibration in a structure. Two coupled, intersecting ultrasonic horns 11, 12 are driven in their respective longitudinal dimensions to excite a respective transverse



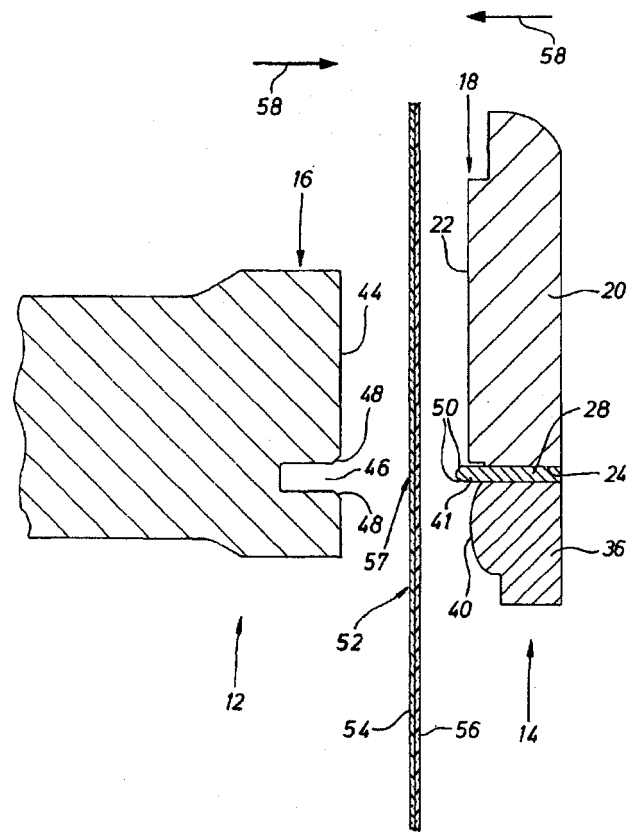
vibration in the other horn. The point at which the horns are joined is thus driven in a composite vibration, which can be transmitted to another member (not shown in the figure) or used to perform an operation such as ultrasonic welding.—IMH

6,811,631

43.40.Sk DEVICE AND METHOD FOR WELDING THIN-WALLED WORK PIECES USING ULTRASOUND

Hartmut Möglich, assignor to Herrmann Ultraschalltechnik
GmbH & Company KG
2 November 2004 (Class 156/73.1); filed in Germany
30 March 2002

This is an ultrasonic welding tool that secures a portion of thin work



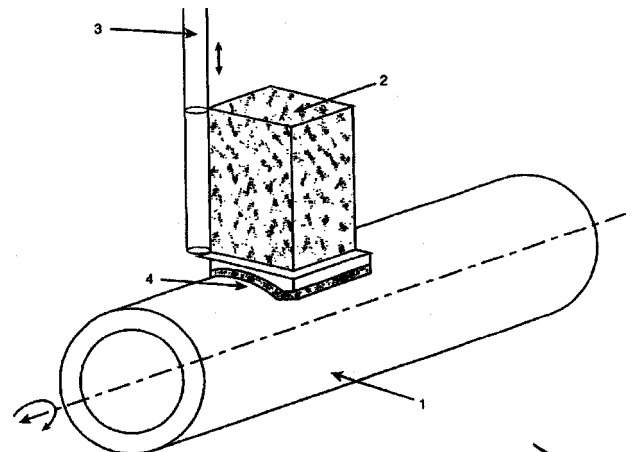
pieces 54, 56 within a mating projection 41 and recess 46.—IMH

6,813,950

43.40.Sk PHASED ARRAY ULTRASONIC NDT SYSTEM FOR TUBES AND PIPES

David Glascock *et al.*, assignors to R/D Tech Incorporated
9 November 2004 (Class 73/633); filed 25 July 2002

This is a system for nondestructive testing of solid tubes. The system includes a phased ultrasonic array 2 having a curved face to match tube 4



being tested. Defects in 3-D (longitudinal, transverse, oblique) are detected by refraction of incident waves in a transmit/receive configuration of the array.—IMH

6,607,359

43.40.Tm APPARATUS FOR PASSIVE DAMPING OF FLEXURAL BLADE VIBRATION IN TURBO-MACHINERY

Andreas H. von Flotow, assignor to Hood Technology Corporation
19 August 2003 (Class 416/229 R); filed 31 January 2002

A scheme to reduce high-order mode vibration in turbine blades consists of a pocket machined into the blade near the tip. A metal insert in this pocket forms the aerodynamic surface of the blade. In response to blade vibration, frictional losses induced by relative motion of the metal insert provide damping.—KPS

6,796,408

43.40.Tm METHOD FOR VIBRATION DAMPING USING SUPERELASTIC ALLOYS

Yang Sherwin and Donald G. Ulmer, assignors to The Boeing Company
28 September 2004 (Class 188/378); filed 13 September 2002

Superelastic, or shape memory alloys, can be deformed repeatedly at high strain levels and still return to their original shapes without accumulating classical fatigue damage. As the load on such an alloy is reversed, considerable energy dissipation occurs as the result of the fully reversible martensitic phase transformation in the alloy. The present patent addresses the use of shape memory alloys for damping of vibrations of such components as turbine blades. Structural dynamic analyses are to be carried out in order to identify areas of high vibratory stress and thermal analyses are to be performed in order to determine the approximate maximum operating temperatures in these areas. The idea is then to dispose a suitably selected alloy in these areas.—EEU

6,799,642

43.40.Tm VIBRATION DAMPER BETWEEN TWO COMPONENTS

Gunter Wolf *et al.*, assignors to Andreas Stihl AG & Company
5 October 2004 (Class 173/162.2); filed in Germany
7 February 2001

This patent pertains to a spring-type isolation system to be placed between the housing of a motor-driven chain saw and a handle thereof. In order to permit the saw to be guided by its operator in the event that the spring breaks, a restraint device is provided in mechanical parallel with the spring, but arranged so that it normally does not short-circuit the spring.—EEU

6,802,405

43.40.Tm FRICTION VIBRATION DAMPER

Carl L. Barcock and Geoffrey R. Tomlinson, assignors to Rolls-Royce plc
12 October 2004 (Class 188/268); filed in the United Kingdom
3 March 2001

This damper is intended to be attached to an engine or machining tool whose vibrations in the vicinity of 10 Hz are to be suppressed. The damper consists of one or more chambers whose volumes are about 95% filled with small particles, which may be spherical with diameters of at least two discrete sizes in the range of 0.1 to 5.0 mm, or elongated, or disc shaped. The chambers are shaped or provided with baffles so that the particles are prevented from moving in convection like flow patterns.—EEU

6,805,053

43.40.Tm METHOD AND DEVICE FOR SUPPRESSING VIBRATIONS IN A PRINTING PRESS

Christopher Berti *et al.*, assignors to Heidelberger Druckmaschinen AG
19 October 2004 (Class 101/480); filed in Germany
20 November 2001

This patent in essence describes single-degree-of-freedom dynamic absorbers whose natural frequencies are automatically tuned to the rotational speeds of selected rollers. A typical torsional absorber consists of a disc mounted at the end of a cantilevered shaft, whose effective length is varied by means of an actuator. The actuator is activated via a controller signal, determined on the basis of the measured printing press rotational speed.—EEU

6,807,863

43.40.Tm VIBRATION PICKUP COMPRISING A PRESSURE SLEEVE

Hartmut Brammer, assignor to Robert Bosch GmbH
26 October 2004 (Class 73/702); filed in Germany
15 December 1999

This patent describes a configuration of an engine knock sensor that can be mounted so as to make secure contact with the vibrating engine. The sensor elements are essentially ring shaped and disposed around a central tubular sleeve, through which a mounting bolt can be placed. The sensor's surface that is to be in contact with the engine has a concave shape to ensure good contact with the engine surface.—EEU

6,814,548

43.40.Tm MULTIPLE PISTON ENGINE WITH VIBRATION REDUCING PROPERTIES

Michel Sagnet, assignor to Peugeot Citroen Automobiles S.A.
9 November 2004 (Class 417/273); filed in France
17 December 1999

This patent relates not to engines and not primarily to vibration reduction, but generally to multipiston pumps and compressors and to their noise reduction. Whereas prior designs achieve noise reduction by arranging the cylinders with angular offsets between their operating cycles, the present patent describes cylinder arrangements that result in different fundamental frequencies for the different cylinders, thus spreading the spectrum of the total generated noise. Different cylinder natural frequencies are achieved by arranging the different pistons to displace different volumes.—EEU

6,806,667

43.40.Vn CONTROL UNIT AND CONTROL METHOD FOR CONTROLLING VIBRATION OF AN ELECTRIC VEHICLE

Shoichi Sasaki and Yoshiaki Ito, assignors to Toyota Jidosha Kabushiki Kaisha
19 October 2004 (Class 318/432); filed 23 May 2003

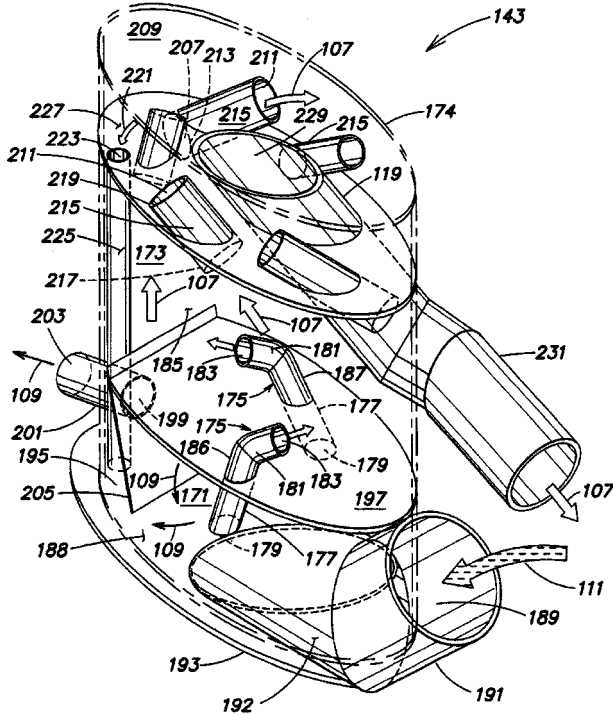
This patent is asserted to be applicable to hybrid vehicles that are driven by a combination of an internal combustion engine and an electric motor. The speed of the electric motor and the speeds of the drive wheel are sensed and the corresponding signals are fed to a controller. The latter provides a torque compensation signal to the electric motor. Several control algorithms are presented, said to require a relatively small amount of computation.—EEU

6,591,939

43.50.Gf MARINE ENGINE SILENCER

Joseph I. Smullin and Matthew E. Denis, assignors to Smullin Corporation
15 July 2003 (Class 181/260); filed 27 April 2001

A muffler for a marine engine includes a chamber that separates the exhaust gas from the cooling water with which it has been mixed. The



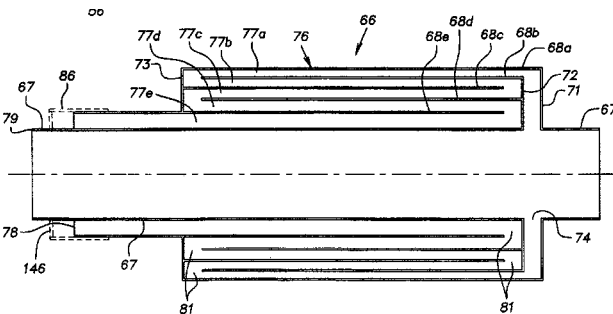
gas/water mixture enters through pipe 189 and flows upwards through tubes 175 which swirl the mixture within the separation chamber. The water mixture exits through 109 and the gas through 107.—KPS

6,595,319

43.50.Gf MUFFLER

Ronald G. Huff, Westlake, Ohio
22 July 2003 (Class 181/250); filed 30 October 2001

An automobile muffler is described which consists of straight pipe 67 and a side branch resonator. The half-wavelength resonator is open at both



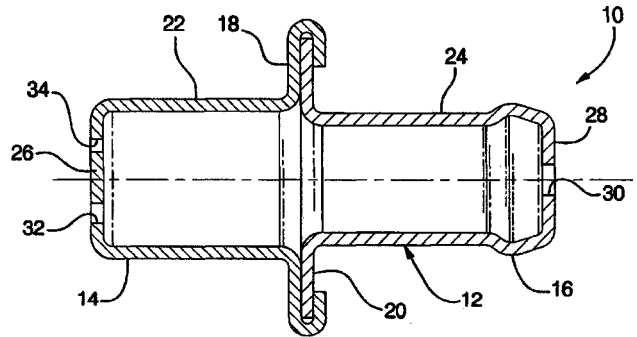
ends. In order to achieve a compact arrangement, the side branch length is folded back on itself. Several alternative geometrical arrangements are described.—KPS

6,619,276

43.50.Gf POSITIVE CRANKCASE VENTILATION ORIFICE MUFFLER

Alan S. Miller and David Thomas Hanner, assignors to General Motors Corporation
16 September 2003 (Class 123/572); filed 28 August 2002

Reduction of noise generated by flow through a positive crankcase ventilation system (PCV) of an automobile is described. The design consists of a cylindrical body 24 with a single output orifice 30 and two inlet orifices



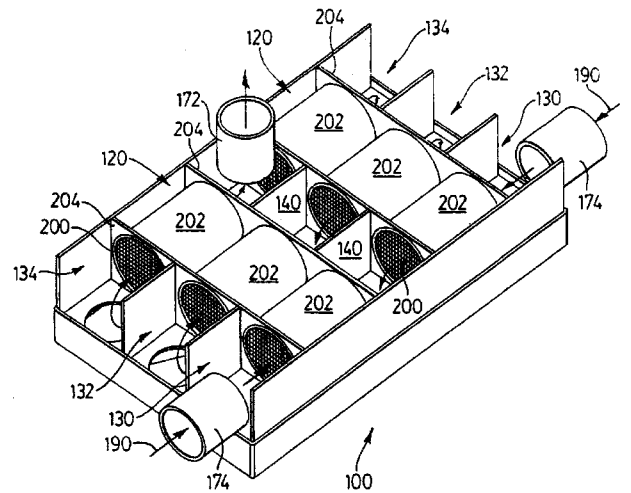
32 and 34. The two inlet orifices have slightly different diameters designed to create "destructive wave interferences" resulting in a 10-dB reduction in sound level at high frequencies.—KPS

6,622,482

43.50.Gf COMBINED CATALYTIC MUFFLER

Glenn Knight and Nils Rodeblad, assignors to Environmental Control Corporation
23 September 2003 (Class 60/299); filed 27 June 2001

A muffler intended for small engines such as those found in lawnmow



ers incorporates a catalytic converter. Arrangements based on cylindrical (see figure) and rectangular geometries are described.—KPS

6,644,061

43.50.Gf SOUND MUFFLING DEVICE FOR AN AIR CONDITIONING UNIT

Joseph W. Woolfson, Brooklyn, New York
11 November 2003 (Class 62/296); filed 2 August 2002

A device for reducing the sound of water droplets falling onto the upper surface of a window or wall-mounted air conditioning unit consists of several screens mounted parallel to one another and located just above the

upper surface of the unit. The first, upper, screen has a coarse mesh, the one below it has finer mesh, and the last one is finer still. Materials such as plastic, metal wire, and nylon screen are recommended and various mounting arrangements are described.—KPS

6,644,437

43.50.Gf VEHICLE EXHAUST WITH LENGTH-EQUALIZING MUFFLER

Alan William Hayman, assignor to General Motors Corporation
11 November 2003 (Class 181/268); filed 2 August 2002

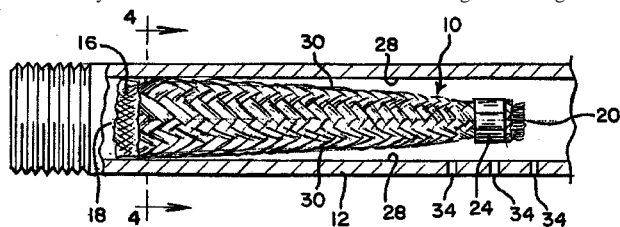
Transverse mounted V-6 and V-8 engines may have exhaust systems with dual exhaust pipes leading from front and rear cylinder banks of the engine to a single muffler. The resultant unequal lengths of pipe lead to a "raspy" exhaust sound quality. A muffler is designed such that it has two internal inlet pipes that have lengths differing by an amount equal to the length difference between the exhaust pipes from the engine, thus achieving approximately overall equal lengths. An additional feature of this muffler design is a valve that is activated by the exhaust pressure, thus achieving dual flow paths and raising the back pressure at low exhaust flow rates.—KPS

6,648,628

43.50.Gf GAS FLOW MUFFLING DEVICE

Scott F. Eiklor, Paoli, Indiana
18 November 2003 (Class 431/114); filed 5 March 2001

Sound generated by burning gas in appliances such as gas logs is reduced by inserting device **10** into the inlet of pipe **18** into which the gas flows. This insert consists of a braided metallic element made from bronze surrounded by a braided fibrous element made of fiberglass. The gas flows



through small orifices formed by the braided insert before flowing through the larger orifices **34** where the gas is burned. It is speculated that the reduction in sound is brought about through a change from laminar to turbulent flow induced by the smaller, upstream orifices.—KPS

6,804,513

43.50.Jh NOISE MONITORING SYSTEM

Tsunehiro Okuya *et al.*, assignors to Honda Giken Kogyo Kabushiki Kaisha
12 October 2004 (Class 455/423); filed in Japan 20 December 2000

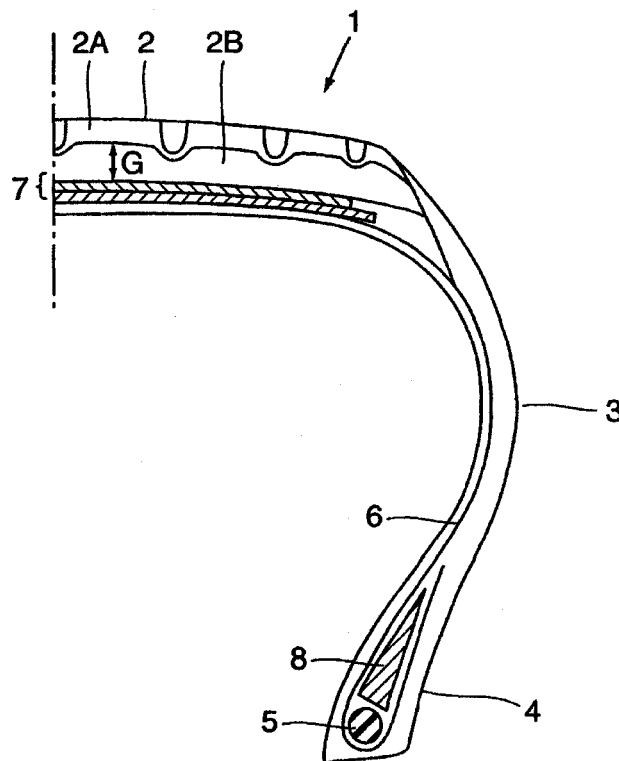
The inventors want to monitor vehicle noise and transmit it via radio when it exceeds a threshold. The whole concept is obvious "to those skilled in the art."—MK

6,626,216

43.50.Lj PNEUMATIC TIRE HAVING HOLLOW PARTICLES IN BASE RUBBER

Akira Minagoshi, assignor to Nokia Corporation
30 September 2003 (Class 152/209.4); filed in Japan 29 June 2000

A method to reduce noise inside a vehicle due to tire-road interaction is described. A tire consists of cap rubber **2A**, which contacts the road surface, attached to a layer of base rubber **2B**. Hollow particles, of diameter



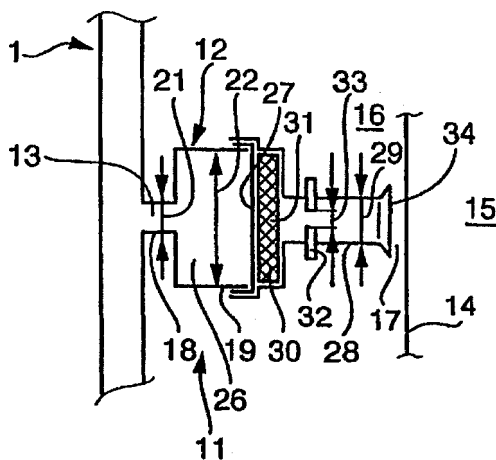
40–200 μm , are mixed in the base rubber and have a volume fraction of 5% to 20%. Details regarding the effects of material selection, particle diameter, and volume fraction on road handling ability and road noise are provided.—KPS

6,600,408

43.50.Lj SOUND TRANSMISSION DEVICE FOR A MOTOR VEHICLE

Norbert Walter *et al.*, assignors to Mahle Filtersysteme GmbH; Bayerische Motorenwerke Aktiengesellschaft
29 July 2003 (Class 340/384.1); filed in Germany 14 May 1999

A device that is designed to introduce engine noise into the passenger compartment of a car is composed of tube **18** connected to air intake **1** and to chamber **26**, thus forming a Helmholtz resonator. A flexible diaphragm **27**



forms an airtight seal. Appropriate choices of geometrical sizes, as well as damping provided by foam 30, enable suitably high sound levels to be transmitted through the engine firewall with no need for penetrations.—KPS

6,612,106

43.50.Nm SEGMENTED MIXING DEVICE HAVING CHEVRONS FOR EXHAUST NOISE REDUCTION IN JET ENGINES

Ronald L. Balzer, assignor to The Boeing Company
2 September 2003 (Class 60/204); filed 30 March 2001

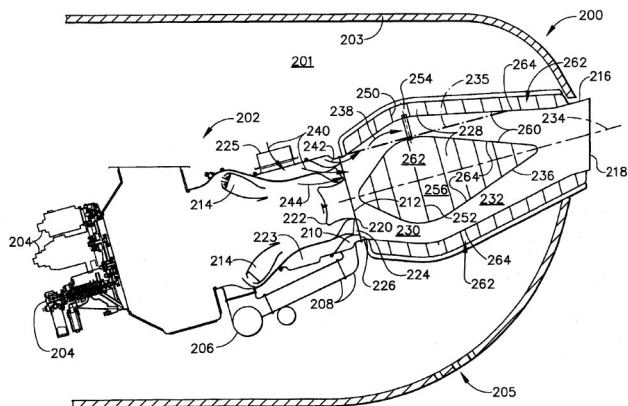
Jet noise from commercial aircraft may be reduced by means of serrated trailing edges on the exhaust nozzles of both the combustor and bypass ducts. In contrast to other designs, this one utilizes rounded rather than sharp serrations. This is said to minimize the generation of unwanted high-frequency noise while improving the desired low-frequency attenuation.—KPS

6,615,576

43.50.Nm TORTUOUS PATH QUIET EXHAUST EDUCTOR SYSTEM

Yogendra Y. Sheoran *et al.*, assignors to Honeywell International Incorporated
9 September 2003 (Class 60/39.5); filed 21 January 2002

Reduction of noise radiated from auxiliary power units commonly found on commercial jet aircraft is achieved via several means. Line-of-



sight from the combustor is blocked, eductor mixing duct 230 and exhaust duct 232 are acoustically treated in the baffled sidewall, and a multi-lobed suppressor is provided at the exhaust exit.—KPS

6,805,633

43.55.Pe GAMING MACHINE WITH AUTOMATIC SOUND LEVEL ADJUSTMENT AND METHOD THEREFOR

Marvin Arthur Hein, Jr. *et al.*, assignors to Bally Gaming, Incorporated
19 October 2004 (Class 463/35); filed 7 August 2002

A system and method are described for automatically adjusting the sound volume of a gaming machine to an appropriate level based on the ambient noise level. The gaming machine may implement an open-loop control algorithm when the machine's sound output is inactive as determined, e.g., by software, or by analyzing samples from a soundboard. A closed-loop control algorithm may be implemented while the gaming machine is active. Alternatively, the machine may implement only an open-loop algorithm that provides volume adjustment only when the machine's sound output is inactive, in which case the adjusted volume is maintained through an active period, then adjusted again during the next inactive period. A saturation back-off feature prevents the sound level of the gaming machine from escalating to a maximum, saturation level.—DRR

6,815,044

43.55.Ti MULTILAYER PANEL WITH AN IMPROVED NOISE DAMPING

Klaus Boock, assignor to Airbus Deutschland GmbH
9 November 2004 (Class 428/172); filed in Germany 15 July 2002

The slots in the core section of this multilayer panel are cut in a curved layout, rather than in typical parallel rows. Also, the depth of the slots varies along their length. In this manner, the panel, which might be used as a cabin divider for a vehicle, has improved transmission loss properties.—CJR

6,796,671

43.58.Wc CONTACT ACTIVATED SOUND AND LIGHT GENERATING NOVELTY FOOD CONTAINERS

Elliot A. Rudell *et al.*, assignors to Rudell Design
28 September 2004 (Class 362/101); filed 21 May 2003

Starting from the fundamental concept of using two electrodes and a conducting comestible, the inventors propose all numbers of different containers that generate light and/or sound.—MK

6,809,634

43.58.Wc BUZZER-DRIVING APPARATUS

Hideaki Uemura and Kenji Wakana, assignors to Sumitomo Wiring Systems, Limited
26 October 2004 (Class 340/384.73); filed in Japan 25 April 2001

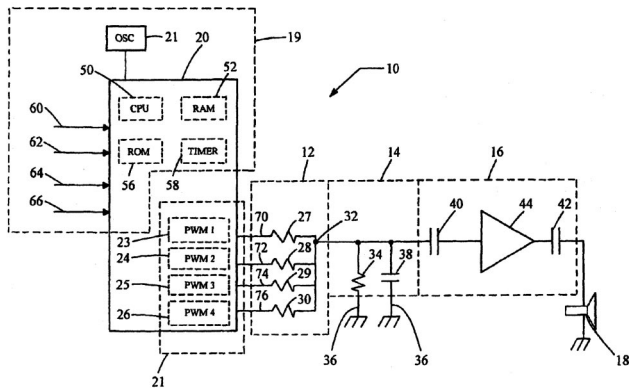
It is well known that any electromechanical device, including buzzers, can be connected to the output of a microcomputer. So, it is not surprising that a single-bit output can control the pitch of a buzzer through its duty cycle, is it? Apparently the authors disagree.—MK

6,798,369

43.58.Wc PRECISION, WIDE BAND PULSE WIDTH MODULATOR FOR DIGITAL TO ANALOG CONVERSION

Charles F. Weber, assignor to Visteon Global Technologies, Incorporated
28 September 2004 (Class 341/144); filed 8 August 2003

Designed for the automotive electronics market, the sound conversion circuit has pulse width modulators 23–26 (internal to microprocessor 19).



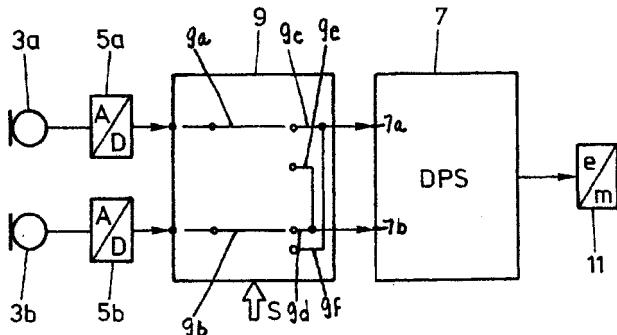
These are summed at node 32 and simply filtered as shown. Simple, crude, but effective.—MK

6,813,363

43.66.Ts PROCEDURE FOR SETTING A HEARING AID, AND HEARING AID

Andi Vonlanthen, assignor to Phonak AG
2 November 2004 (Class 381/313); filed in the World IPO
14 October 1999

Each microphone of a two-microphone directional system used in a hearing aid is normally designated to be a front or rear microphone. The outputs of the microphones connecting to the signal processor in the directional system must be configured to have the axis of directionality pointing forward whether the hearing aid is worn in the left or right ear. A switching



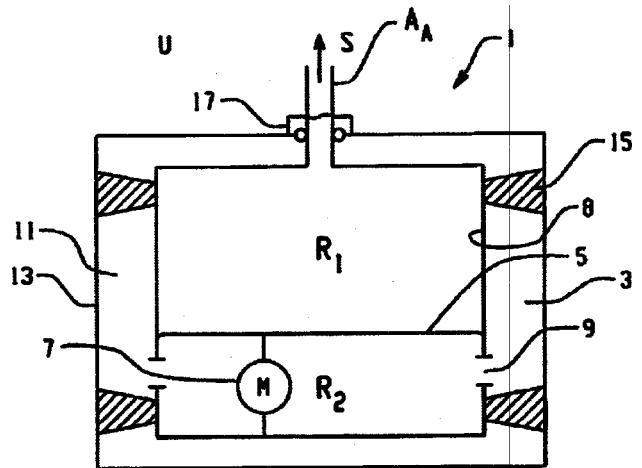
network is implemented under fitting software control to adjust the effective position of the microphones for use in either left or right hearing aids, thus enabling a single universal microphone configuration to be manufactured.—DAP

6,813,364

43.66.Ts ELECTRIC/ACOUSTIC TRANSDUCER MODULE, IN-EAR HEARING AID AND METHOD FOR MANUFACTURING AN IN-EAR HEARING AID

Andi Vonlanthen, assignor to Phonak AG
2 November 2004 (Class 381/322); filed in the World IPO
16 June 1999

A resilient transducer isolator system is described for keeping the loudspeaker floating, preventing it from touching the hearing aid housing. Low-frequency response is enhanced by additional ports to the isolator gap from behind the diaphragm of the loudspeaker. The system permits automatic



insertion of the transducer module into the hearing aid housing. An acoustically transparent membrane prevents debris from obstructing the hearing aid output and also helps to acoustically seal the transducer output from the rest of the hearing aid.—DAP

6,815,434

43.66.Vt METHODS FOR TREATING HEARING LOSS

Jonathan Kil and Eric D. Lynch, assignors to Sound Pharmaceuticals Incorporated
9 November 2004 (Class 514/58); filed 3 January 2003

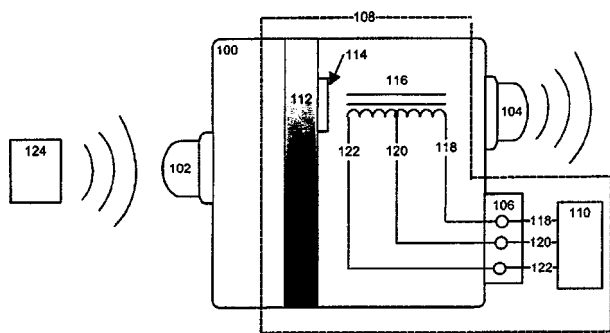
Each of the methods described in the patent document consists of the step of administering to a subject an amount of an otoprotectant composition reportedly able to ameliorate hearing loss. In some embodiments, the otoprotective composition consists of a pharmaceutically effective amount of at least one glutathione peroxidase mimic (i.e., a composition comprising enselen). In some embodiments, among others, the otoprotective composition consists of a supposedly pharmaceutically effective amount of (a) at least one glutathione peroxidase mimic and (b) at least one glutathione or glutathione precursor (e.g., a composition comprising allopurinol and N-acetyl-cysteine).—DRR

6,768,803

43.72.Ew METHOD AND APPARATUS FOR SELECTIVE ACOUSTIC SIGNAL FILTERING

Gilles Duhamel, assignor to Sonomax Hearing Healthcare Incorporated
27 July 2004 (Class 381/372); filed 20 March 2000

Diaphragm 112 is excited by sound entering enclosure 100 through port 102 from sound source 124. Magnet 114 is attached to diaphragm 112.



Passive and/or active components in load **110** and connected through electrical port **106** can shape the response of induction coil **116** which in turn couples to **112** through **114**. Thus, the sound emanating from acoustic port **104** can be further modified from that of the original. Since items **100**, **102**, **104**, and **116** are present in many hearing aids, the patent describes an additional measure of frequency response modification that can be added to the aid.—NAS

6,813,490

43.72.Ew MOBILE STATION WITH AUDIO SIGNAL ADAPTATION TO HEARING CHARACTERISTICS OF THE USER

Heikki Lang *et al.*, assignors to Nokia Corporation
2 November 2004 (Class 455/414.1); filed 17 December 1999

The real novelty proposed in this patent is a system that is promised to custom enhance a speech signal for the benefit of a user whose audiometric profile (or some set of parameters specifying individual hearing ability) is known. The authors propose to make speech more intelligible by transposing the spectrum of sounds to a frequency range that can be heard by the user, while keeping the essential frequency relations among formants intact. To those well-versed in the art, such a proposal is recognizable as a likely pipe dream and, as is typical of such, the patent fails to provide any implementable (or even meaningful) indication of just how such a miraculous speech frequency transposition is to be accomplished. The patent spends a great deal of space detailing how such a system could be integrated into a mobile user station, relegating the all-important modification of the speech signal to a black box in the diagrams. The best that can be said here is that certainly the goal of the patented novelty is a laudable one. It is surprising that an unattained goal has been patented.—SAF

6,810,378

43.72.Ja METHOD AND APPARATUS FOR CONTROLLING A SPEECH SYNTHESIS SYSTEM TO PROVIDE MULTIPLE STYLES OF SPEECH

Gregory P. Kochanski and Chi-Lin Shih, assignors to Lucent Technologies Incorporated
26 October 2004 (Class 704/258); filed 24 September 2001

Explicitly based on the methods employed by human mimics of the speech of other humans, this patent describes a system using a database of templates controlling patterns of fundamental frequency, amplitude characteristics, etc. The method selects templates matching the characteristics of a particular speaker and uses these templates to generate synthesized speech mimicking the speaker. Some of the relevant information is incorporated into the patent by reference to another patent application, which at present does not appear in patent databases.—HRJ

6,813,604

43.72.Ja METHODS AND APPARATUS FOR SPEAKER SPECIFIC DURATIONAL ADAPTATION

Chi-Lin Shih and Jan Pieter Hendrik van Santen, assignors to Lucent Technologies Incorporated
2 November 2004 (Class 704/260); filed 13 November 2000

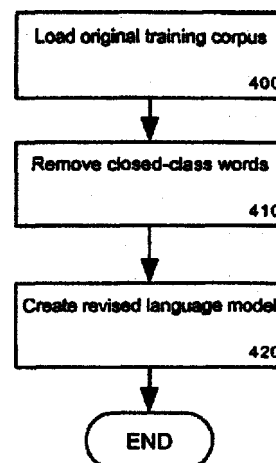
The durational patterns of different people's speech can vary in non-uniform ways—different speakers will have different durations for different speech sounds, and vary these durations to a different extent depending on speech rate, sentence structure, etc. This patent describes a method for determining an individual's durational characteristics for different sounds, modeling those characteristics, and adapting a text-to-speech system's durational patterns to follow an individual speaker's model.—HRJ

6,772,116

43.72.Ne METHOD OF DECODING TELEGRAPHIC SPEECH

James R. Lewis, assignor to International Business Machines Corporation
3 August 2004 (Class 704/231); filed 27 March 2001

Decoding telegraphic speech such as newspaper headlines requires different speech recognition techniques than conventional methods used to recognize speech that contains closed class words such as "a" or "the."



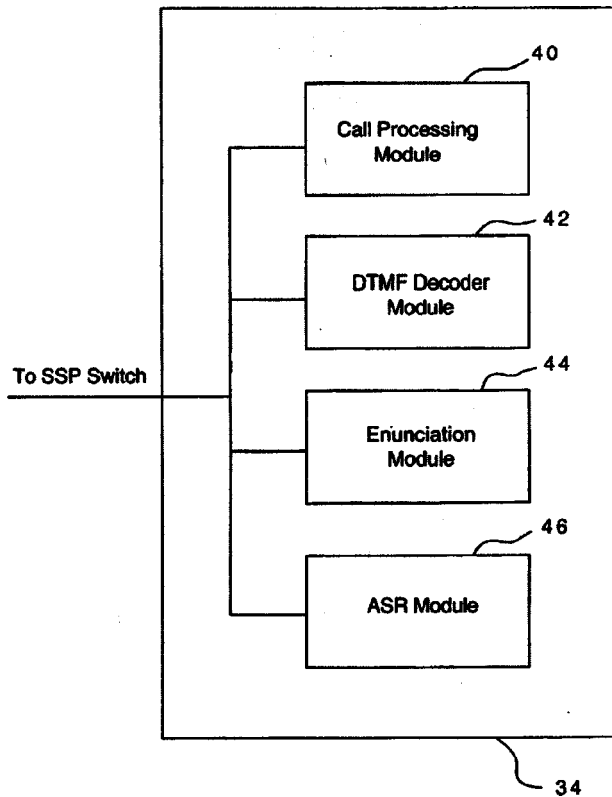
Described is a method for selectively applying telegraphic language models to user-spoken utterances in order to improve decoding accuracy.—DAP

6,778,640

43.72.Ne NETWORK AND METHOD FOR PROVIDING A USER INTERFACE FOR A SIMULTANEOUS RING TELECOMMUNICATIONS SERVICE WITH AUTOMATIC SPEECH RECOGNITION CAPABILITY

Zeeman Zhang and Joseph Knoerle, assignors to BellSouth Intellectual Property Corporation
17 August 2004 (Class 379/88.01); filed 9 August 2000

A speech recognition system allows telecommunications users to utilize advanced calling features by speaking predetermined voice commands over either landline or wireless telephones rather than having to dial numerical codes. A switch detects triggering events and routes the communications



to an intelligent resource server that contains speech recognition capability and an enunciation module that produces audible prompts and responses.—DAP

6,778,641

43.72.Ne TELEPHONE INDEPENDENT PROVISION OF SPEECH RECOGNITION DURING DIAL TONE AND SUBSEQUENT CALL PROGRESS STATES

Paul Andrew Erb, assignor to Mitel Knowledge Corporation
17 August 2004 (Class 379/88.01); filed in the United Kingdom
13 September 2000

A speech recognition engine resource is allocated automatically during calls from telephone device users with no additional intervention by the user.—DAP

6,785,654

43.72.Ne DISTRIBUTED SPEECH RECOGNITION SYSTEM WITH SPEECH RECOGNITION ENGINES OFFERING MULTIPLE FUNCTIONALITIES

James Cyr *et al.*, assignors to Dictaphone Corporation
31 August 2004 (Class 704/270.1); filed 30 November 2001

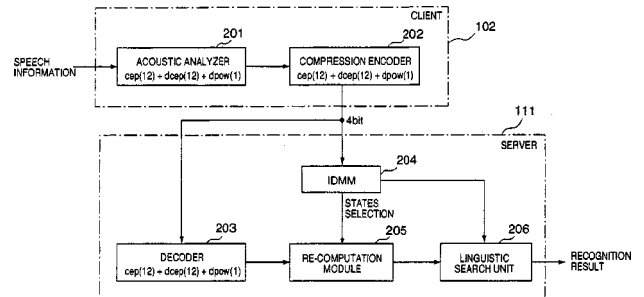
Processed speech inputs are stored until they are forwarded to one of several speech recognition engines. There are several servers performing different functions in each speech recognition engine, each of which can be activated when needed or deactivated when not in use. As more users are added, the distributed system is said to provide more optimal performance as the load increases as compared to a central processing facility.—DAP

6,813,606

43.72.Ne CLIENT-SERVER SPEECH PROCESSING SYSTEM, APPARATUS, METHOD, AND STORAGE MEDIUM

Teruhiko Ueyama *et al.*, assignors to Canon Kabushiki Kaisha
2 November 2004 (Class 704/270.1); filed in Japan 24 May 2000

To minimize the amount of data transfer between a client and a server, speech parameters are compression-encoded before transfer to the server.



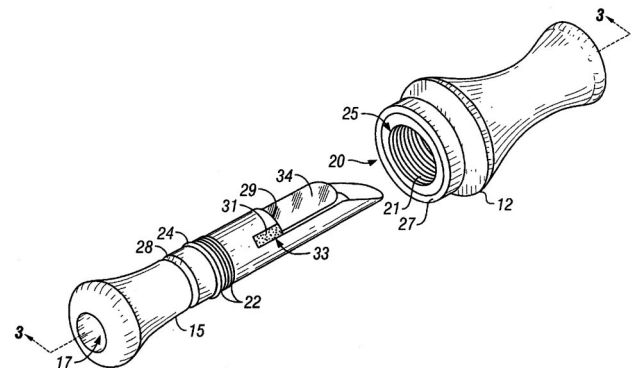
After 25-dimensional speech parameters are produced by an acoustic analysis, they are quantized and encoded into compressed parameters at four bits per dimension per frame.—DAP

6,435,933

43.75.Ef GAME CALL APPARATUS

Robert E. Browne, Sharon, Tennessee
20 August 2002 (Class 446/207); filed 20 November 2001

This well written disclosure illustrates how a capped reed (ala Krumhorn) can be used as a duck or goose call. Standard modern woodworking



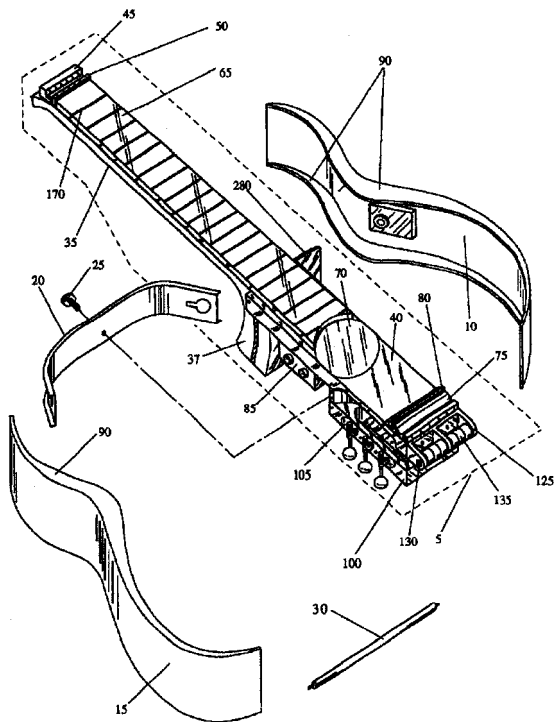
construction is used throughout: o-rings, acrylic bodies, and plastic reeds.—MK

6,791,022

43.75.Gh STRINGED MUSICAL INSTRUMENT

Philip Shepard Green, Palo Alto, California
14 September 2004 (Class 84/731); filed 25 February 2002

Travelling with a guitar can be inconvenient. This invention is a portable, decomposable electric guitar. As shown in the figure, note how the body is shaped like an acoustic body. The support arm 20 allows the sides 10 and 15 to be connected. The connecting rod 30 is not connected to a



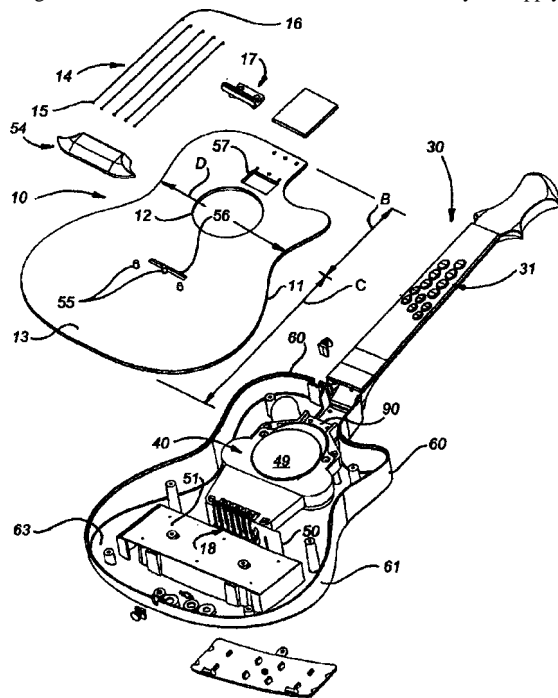
piston but adds stiffness. The fretboard and string tensioners are affixed via 280. Lacking a body resonance, the sound must be electronically processed.—MK

6,800,797

43.75.Gh METHOD AND APPARATUS FOR PRODUCING ACOUSTICAL GUITAR SOUNDS USING AN ELECTRIC GUITAR

Henry Martin Steiger III, Phoenix, Arizona
5 October 2004 (Class 84/267); filed 15 February 2002

Once again, we see an electric guitar trying to masquerade as an acoustic guitar. The inventor wants to use the hollow body to supply body

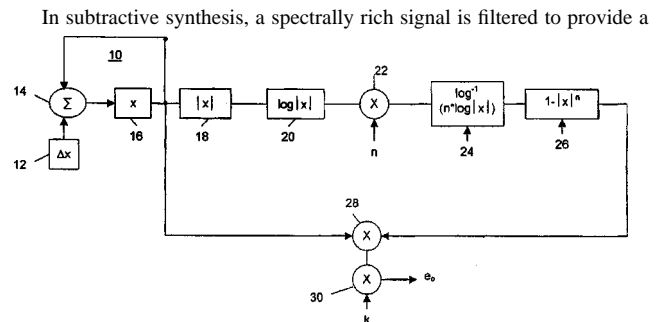


resonance but it's filled with the mechanical strummer 18 and loudspeakers 40. Note the use of a keyboard 31 rather than frets.—MK

6,806,413

43.75.Wx OSCILLATOR PROVIDING WAVEFORM HAVING DYNAMICALLY CONTINUOUSLY VARIABLE WAVESHAPE

Robert Chidlaw and Ralph Muha, assignors to Young Chang Akki Company, Limited
19 October 2004 (Class 84/660); filed 31 July 2002



palette of interesting timbres. The proposed generator is $y = x \times (1 - x^n)$. As shown in the figure, this can be done via table lookup.—MK

6,809,673

43.75.Yy MULTI-CHANNEL CIRCUIT WITH CURRENT STEERING DIGITAL TO ANALOGUE CONVERTERS WITH MINIMIZED CROSSTALK

Anthony Scanlan and John Patrick Purcell, assignors to Analog Devices, Incorporated
26 October 2004 (Class 341/144); filed 10 October 2002

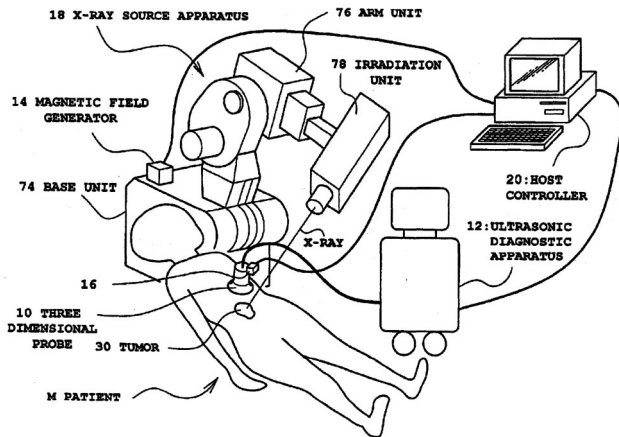
To isolate the current switches in a digital-to-analog converter, it is "standard operating procedure" to provide external pins for decoupling capacitors (providing on-chip capacitors would be prohibitively expensive). The inventors present an alternative—the use of cascode switches connected to the power rails. The gate capacitance will provide some feedthrough but this is not analyzed or discussed.—MK

6,796,943

43.80.Qf ULTRASONIC MEDICAL SYSTEM

Takashi Mochizuki, assignor to Aloka Company, Limited
28 September 2004 (Class 600/437); filed in Japan 27 March 2002

The intent of this ultrasonic system is to locate tumors so that they can be treated by radiation in an accurate fashion. The ultrasonic diagnostic apparatus 12 sends echo data to a host controller 20. In the host controller, a probe coordinate operation unit computes the coordinates of 3-D probe 10



using x-ray source 18 as a reference position. A tissue coordinate operation unit computes the coordinates of tumor 30 using the probe as a reference origin. The tumor coordinates are relayed to the x-ray irradiation unit 78, so that the tumor is targeted and irradiated.—DRR

6,802,813

43.80.Qf METHODS AND APPARATUS FOR MONITORING AND QUANTIFYING THE MOVEMENT OF FLUID

Ernest G. Schutt, San Diego, California
12 October 2004 (Class 600/454); filed 23 May 2003

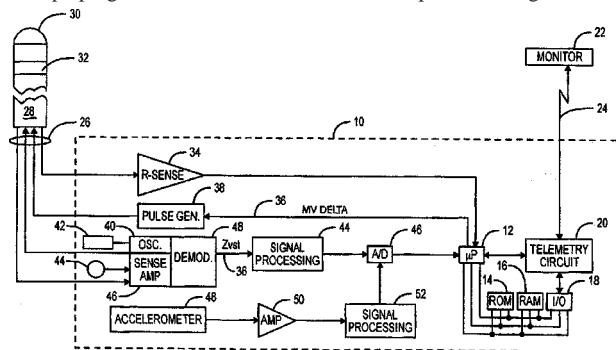
A methodology is ostensibly to be provided for monitoring and quantifying the movement of fluid in a target region. Generally, an imaging agent is introduced into a target region through fluid flow. The imaging agent in the target region is then disrupted using appropriate methods, such as the application of ultrasonic energy. As fluid flow brings undisrupted imaging agent into the target region, the rate of accumulation is monitored and quantified, thereby providing the exchange rate and flow rate of the fluid in the target region. This method is said to be useful for medical applications such as determining the flow rate of blood in an organ or tissue.—DRR

6,810,287

43.80.Qf IMPLANTABLE CARDIAC DISEASE MANAGEMENT DEVICE WITH TRIGGER-STORED POLYSOMNOGRAM AND PHONOCARDIOGRAM

Qingsheng Zhu *et al.*, assignors to Cardiac Pacemakers, Incorporated
26 October 2004 (Class 607/17); filed 3 December 2001

The device consists of a body-implantable cardiac rhythm management device (CRMD) that has the capability of recording polysomnogram (PSG) data and/or phonocardiogram (i.e., heart sound) data upon the detection of preprogrammed conditions or events. The polysomnograms can be



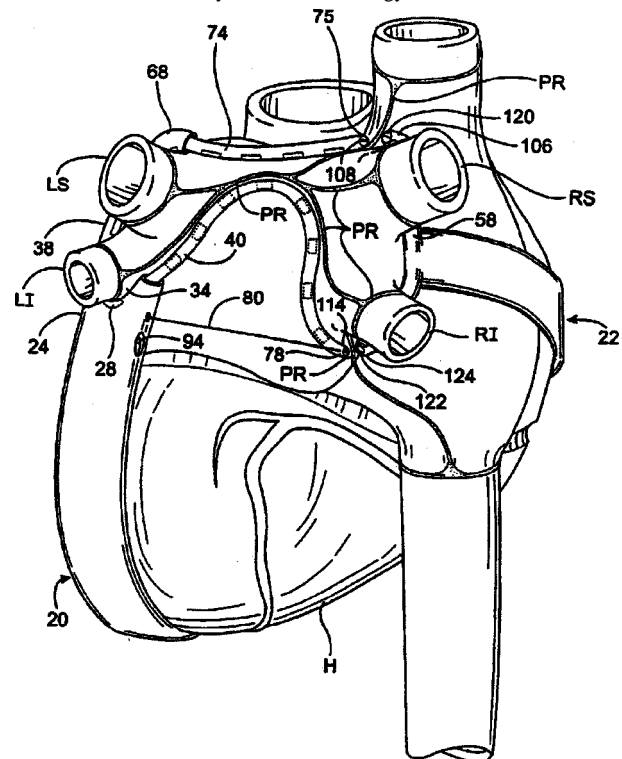
obtained from a sensor of the type used in minute ventilation-based, rate-adaptive pacemakers. The phonocardiogram data is obtained from an accelerometer transducer or other type of sensor capable of detecting heart sounds.—DRR

6,805,128

43.80.Sh APPARATUS AND METHOD FOR ABLATING TISSUE

Benjamin Pless *et al.*, assignors to Epicor Medical, Incorporated
19 October 2004 (Class 128/898); filed 12 July 2000

We have here a system that is intended for the diagnosis and treatment of electrophysiological diseases, more specifically, for epicardial mapping and ablation for the treatment of arterial fibrillation. A control system alters one or more characteristics of an ablating element to ablate tissue. In one embodiment, the control system delivers energy nearer to the tissue surface



by changing the frequency or power. In another embodiment, the ablating element delivers focused ultrasound that is focused in at least one dimension. The ablating device may also have a number of ablating elements with different characteristics, such as focal length.—DRR

6,814,702

43.80.Sh APPARATUS FOR MEASURING HEMODYNAMIC PARAMETERS

Richard T. Redano, assignor to Neutrino Development Corporation
9 November 2004 (Class 600/454); filed 27 April 2002

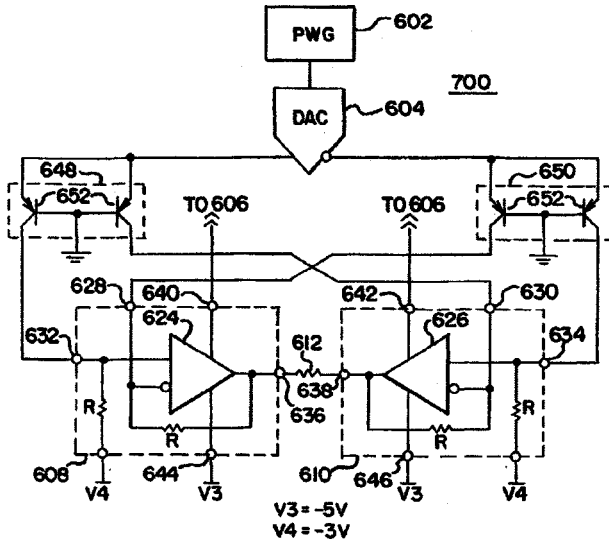
Apparently intended as an alternative to Viagra, this device is aimed at treating vasculogenic erectile dysfunction that results from inadequate blood flow into the penis. The first step of the method is coupling an ultrasound source to the target. The ultrasound source is coupled to a lesion-free region of the penile outer surface. The second step of the method is transmitting ultrasound energy into the corpora cavernosum at a sufficient frequency and intensity to increase hemodynamic flow. The frequency applied is a function of the depth of desired penetration into the corpora cavernosum.—DRR

6,808,494

43.80.Vj TRANSMIT CIRCUIT FOR IMAGING WITH ULTRASOUND

Lazar A. Shifrin, assignor to Siemens Medical Solutions USA, Incorporated
26 October 2004 (Class 600/437); filed 10 February 2003

This circuit consists of a digital-to-analog converter with differential



outputs, current splitters, difference amplifiers, and a push-pull output amplifier.—RCW

6,808,495

43.80.Vj ULTRASONIC CELLULAR TISSUE SCREENING TOOL

Kevin M. Kelly *et al.*, assignors to Sonocine, Incorporated
26 October 2004 (Class 600/443); filed 23 December 2002

An ultrasound probe is moved across a region of tissue to produce sequential, closely spaced images. The linear position and angular orienta-

tion of the probe are monitored and controlled during the movement. The images and the position and orientation data are recorded for manipulation and display.—RCW

6,811,766

43.80.Vj ULTRASOUND IMAGING WITH CONTRAST AGENT TARGETED TO MICROVASCULATURE AND A VASODILATOR DRUG

Morten Eriksen *et al.*, assignors to Amersham Health AS
2 November 2004 (Class 424/9.52); filed in the United Kingdom
21 October 1997

A combination of an ultrasound contrast agent and a vasodilator drug is used. The contrast agent accumulates in concentrations related to the regional rate of tissue perfusion. The vasodilator drug enhances the difference between normally perfused and underperfused tissue.—RCW

6,814,701

43.80.Vj METHOD AND APPARATUS FOR ULTRASOUND DIAGNOSTIC IMAGING

Tadashi Tamura, assignor to Aloka Company, Limited
9 November 2004 (Class 600/443); filed 16 April 2003

Speckle noise in ultrasound images is reduced by forming overlapping sector images from different apexes and combining the images to produce a compound image.—RCW

LETTERS TO THE EDITOR

This Letters section is for publishing (a) brief acoustical research or applied acoustical reports, (b) comments on articles or letters previously published in this Journal, and (c) a reply by the article author to criticism by the Letter author in (b). Extensive reports should be submitted as articles, not in a letter series. Letters are peer-reviewed on the same basis as articles, but usually require less review time before acceptance. Letters cannot exceed four printed pages (approximately 3000–4000 words) including figures, tables, references, and a required abstract of about 100 words.

An artificial neural network approach for predicting architectural speech security (L)

Jingfeng Xu

School of Architecture, Design Science and Planning, University of Sydney, NSW 2006, Australia

John S. Bradley and Bradford N. Gover

Institute for Research in Construction, National Research Council, 1200 Montreal Road, Ottawa, Ontario K1A 0R6, Canada

(Received 26 August 2004; revised 6 January 2005; accepted 6 January 2005)

Signal-to-noise type measures have been developed for predicting architectural speech privacy and speech security, which is required to accurately rate the probability of a listener outside a room being able to overhear conversations from within the room. However, these measures may not be ideal for speech security situations. In the present work, an approach that uses the artificial neural networks to directly represent the functional relationship between the octave band (250 Hz–8 kHz) S/N ratios and the speech intelligibility score and security thresholds has been investigated. The artificial neural network approach provides a direct and accurate method for predicting the speech intelligibility score and security thresholds. © 2005 Acoustical Society of America.

[DOI: 10.1121/1.1862092]

PACS numbers: 43.55.Hy, 43.71.Gv, 43.60.Lq [NX]

Pages: 1709–1712

I. INTRODUCTION

Speech security can be described on three different levels. The first level would be when only a very small percentage or none of the overheard words is intelligible. The second level is when no words are intelligible and it is often still possible to recognize the cadence or rhythm of the speech. Finally, the highest level of speech security would be when all speech sounds from the adjacent space are completely inaudible.¹ Speech privacy and speech security have been related to signal-to-noise (S/N) type measures, where the signal is the speech from the adjacent space. The simplest measure is the difference of A-weighted speech and noise levels. More sophisticated measures such as the Articulation Index (AI) (Ref. 2) and its more recent replacement, the Speech Intelligibility Index (SII),³ are known to be better related to speech intelligibility within rooms. However, Gover and Bradley¹ have shown that SII and AI cannot be used to describe conditions for high levels of speech security which would correspond to acoustical conditions below $SII=0$, where SII is not defined. The difference in A-weighted levels is not limited in this way but it is much less accurately related to intelligibility scores. A more successful measure is the SII-weighted S/N ratio, which is a weighted sum of one-third-octave-band S/N ratios using the same frequency weightings as the SII measure. Although the SII-weighted

S/N ratio predicts the intelligibility score and threshold reasonably well, the S/N loudness ratio provides a more accurate estimation of the thresholds of cadence and audibility.¹

In the present work, an approach that uses the artificial neural networks (ANNs) to directly represent the functional relationship between the octave band (250 Hz–8 kHz) S/N ratios and the speech intelligibility score and security thresholds has been investigated. The objectives of the ANNs are to predict (a) the speech intelligibility score, namely the percentage of words correctly identified by each individual; (b) the intelligibility threshold, namely the percentage of listeners able to correctly identify at least one word; (c) the ca-

TABLE I. Ranges of input variables.

Input variables	Range of input variables for predicting the intelligibility scores and the intelligibility threshold		Range of input variables for predicting the cadence threshold and the audibility threshold	
	Minimum dB	Maximum dB	Minimum dB	Maximum dB
S/N Ratio at 250 Hz	-34.8	6.2	-34.8	-8.3
S/N Ratio at 500 Hz	-30.2	13.1	-30.2	-2.3
S/N Ratio at 1 kHz	-32.7	1.3	-32.7	-15.7
S/N Ratio at 2 kHz	-28.2	7.1	-28.2	-1.8
S/N Ratio at 4 kHz	-27.4	10.0	-21.8	0.5
S/N Ratio at 8 kHz	-23.2	8.7	-16.6	-6.3

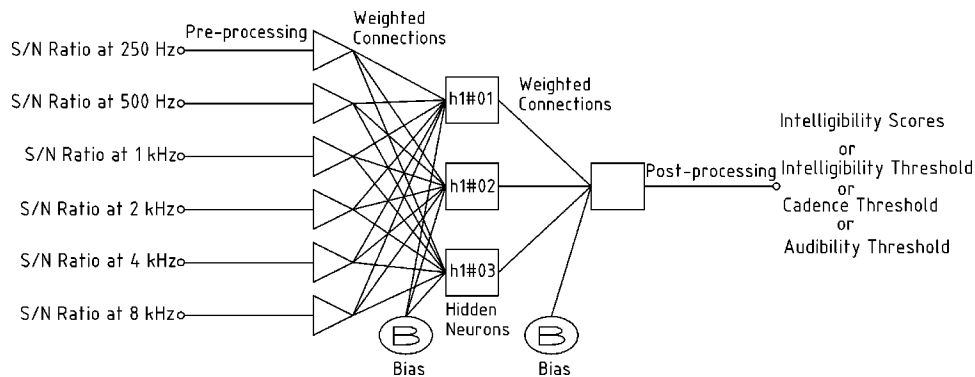


FIG. 1. Illustration of a three-layer feed-forward artificial neural network.

dence threshold, namely the percentage of listeners able to detect the cadence of the speech; and (d) the audibility threshold, namely the percentage of listeners able to hear the presence of the speech.

The ANN approach provides a direct and accurate method for predicting the speech intelligibility score and security thresholds. ANNs are inherently capable of representing nonlinear systems. They learn from historical data and model input–output functional relationships. The history and theory of ANNs, their advantages and shortcomings in applications, and their future utility have been presented elsewhere^{4–6} and will not be presented here. A brief overview of how ANNs operate is presented by Nannariello *et al.*⁷ and Li *et al.*⁸

II. DATABASE FOR ANN MODELS

Two speech intelligibility and security experiments were carried out by Gover and Bradley¹ at the Institute for Research in Construction (IRC), National Research Council Canada (CNRC). In the first intelligibility experiment (intelligibility scores and the intelligibility threshold), 36 subjects each listened to 340 test sentences. Subjects were then divided into two groups, 19 “better” subjects and 17 “worse” subjects, in accordance with their mean intelligibility scores across all 340 sentences. A follow-up experiment intended to determine not only the intelligibility score and the intelligibility threshold but also the thresholds of cadence of the speech and audibility of any speech sounds. In this second experiment the 19 better subjects from the first experiment each listened to 160 sentences. Details of these two experiments are presented by Gover and Bradley.¹

The results of 19 better subjects in the above two experiments constitute the database for the ANN models in the present work. The 9500 [19×(340+160)] individual intelligibility score test cases were used as the database for predicting the intelligibility score. Of the 9500 cases, 7600 were used for training, 950 for verification, and 950 for testing. All the sentences in the first and second experiments, *viz.*, 500 (340+160) sentences (cases), were used as the database for predicting the intelligibility threshold. Of the 500 cases, 310 were used for training, 95 for verification, and 95 for testing. For predicting the cadence threshold and the audibility threshold, the 160 sentences (cases) in the second experiment were used as the database. Of the 160 cases, 128 were used for training, 16 for verification, and 16 for testing.

III. INPUTS AND OUTPUTS OF ANN MODELS

The inputs of the ANN models for all four prediction situations are the octave band (250 Hz–8 kHz) S/N ratios, that is to say, the difference in the transmitted speech level and the background noise level at each octave frequency band from 250 Hz to 8 kHz. The range of the input variables for predicting the intelligibility score and the intelligibility threshold when considering both experiments is different from that for predicting the audibility and the cadence thresholds when accounting for only the second experiment. Table I provides the range of each input variable of the four ANN models. The outputs of the ANN models are the intelligibility score, the intelligibility threshold, the cadence threshold, and the audibility threshold, and are within the range of 0%–100%.

TABLE II. Shift and scale factors to inputs of each ANN model.

Input variables	ANN model for predicting intelligibility scores		ANN model for predicting intelligibility threshold		ANN model for predicting cadence threshold		ANN model for predicting audibility threshold	
	Shift	Scale	Shift	Scale	Shift	Scale	Shift	Scale
S/N Ratio at 250 Hz	0.8478	0.0244	0.8561	0.0251	1.3117	0.0377	1.3117	0.0377
S/N Ratio at 500 Hz	0.6984	0.0231	0.6984	0.0231	1.0840	0.0363	1.0953	0.0362
S/N Ratio at 1 kHz	0.9629	0.0295	0.9748	0.0298	1.9676	0.0605	1.9299	0.0590
S/N Ratio at 2 kHz	0.7990	0.0284	0.7990	0.0284	1.1847	0.0421	1.0704	0.0382
S/N Ratio at 4 kHz	0.7333	0.0268	0.7333	0.0268	0.9791	0.0450	1.0969	0.0503
S/N Ratio at 8 kHz	0.7276	0.0314	0.7254	0.0316	1.6072	0.0968	1.6063	0.0966

TABLE III. Weights and biases to hidden neurons of each ANN model.

	ANN model for predicting intelligibility scores			ANN model for predicting intelligibility threshold			ANN model for predicting cadence threshold			ANN model for predicting audibility threshold		
	h1#01	h1#02	h1#03	h1#01	h1#02	h1#03	h1#01	h1#02	h1#03	h1#01	h1#02	h1#03
Bias	11.5184	3.4952	2.4780	4.7585	-1.0361	-3.6038	-1.3096	-3.4213	0.0453	1.6085	-0.0568	0.8694
S/N Ratio at 250 Hz	1.1366	1.4757	0.7391	2.3090	4.7684	-0.5849	-0.6167	-2.5940	0.9785	3.0270	4.2765	-1.6642
S/N Ratio at 500 Hz	10.5723	2.7610	4.5765	2.9866	-0.8594	-4.8108	-0.7203	-1.2851	0.3082	1.9515	2.4088	-1.6567
S/N Ratio at 1 kHz	-0.2854	4.9302	7.2539	3.1024	-2.6401	-9.3399	-1.8440	-0.1070	0.4341	-1.3097	0.1111	1.1948
S/N Ratio at 2 kHz	8.0602	1.3246	1.8921	2.1874	1.4327	-1.4269	1.0129	-4.0315	2.1161	4.1976	1.1180	-1.7192
S/N Ratio at 4 kHz	2.0982	2.0439	-0.0099	2.4986	0.7054	-0.0970	2.5294	-6.9358	2.9931	6.7447	1.3925	-3.6152
S/N Ratio at 8 kHz	1.1966	-1.0482	-2.9519	1.5709	2.3245	-0.1919	-1.8705	3.6446	-0.4777	-3.5032	-0.2028	2.2250

IV. ARCHITECTURE OF ANN MODELS AND PROCEDURES OF ANN ANALYSIS

Figure 1 illustrates the three-layer feed-forward ANN architecture applied in the present work. One hidden neuron layer, with three hidden neurons, was used. The ANN analysis was undertaken using STATISTICA Neural Networks.⁹ The weights, which were initialized to uniformly distributed random values using the “uniform method,” were adjusted by using “back propagation” and “conjugate gradient” algorithms to minimize the prediction error during the training. Conditions were set within the STATISTICA Neural Networks program by altering the model’s parameters. “Early stopping” and “Weigend weight regularization” techniques were used to control overfitting. Training of ANNs was stopped when the rms error of the verification set could no longer be improved. The test set was used to independently check the performance of the network when an entire network design procedure was completed.

V. MODUS OPERANDI OF ANN MODELS

For the present work, the preprocessing of the inputs involves scaling input values to an appropriate range suitable for use in the ANN. Input values are multiplied by a scale factor, followed by the addition of a shift factor. Table II provides the shift and scale factors to input variables of each ANN model. These factors are automatically generated by the scaling algorithm of the STATISTICA Neural Networks.⁹

Each preprocessed neuron is multiplied by a scalar weight connecting the first-layer neurons to the hidden-layer neurons. At each neuron within the hidden layer, the weighted inputs are summed and bias value is subtracted

from the summed weighted inputs. The resulting value is passed through a nonlinear activation function, in this work, sigmoidal logistic function ($1/(1+e^{-x})$, where x is the resulting value). Table III provides the weights and biases linked to hidden neurons of each ANN model attained by training.

The output value of each hidden neuron is multiplied by the scalar weights for each connection between the hidden layer and the output neuron. The weighted outputs are summed and a bias value is subtracted from the sum to produce a single output value. Table IV provides the weights and biases linked to the output neuron of each ANN model attained by training.

The output is postprocessed by subtracting the shift factor, followed by division by the scale factor. In the present work, the shift factor and scale factor are 0 and 1, respectively, in all four situations as the range of outputs has already been normalized between 0 and 1 in the original database.

Work carried out by Nannariello *et al.*¹⁰ and Xu *et al.*¹¹ presented a method of how to embed an ANN model into a standard spreadsheet. The same procedure can be adopted with those details provided in Table II–Table IV.

VI. RESULTS OF ANN ANALYSIS

Table V presents the statistical analysis results based on the ANN models specified in the above section. In a regression problem, the standard deviations (S.D.) of both the prediction error and original output data are important and they can be related by the S.D. ratio. The S.D. ratio in a regression problem is the ratio of the S.D. of the prediction error to

TABLE IV. Weights and biases to the output neuron of each ANN model.

	ANN model for predicting intelligibility scores	ANN model for predicting intelligibility threshold	ANN model for predicting cadence threshold	ANN model for predicting audibility threshold
	Output (intelligibility scores)	Output (intelligibility threshold)	Output (cadence threshold)	Output (audibility threshold)
Bias	0.0086	-0.6521	-1.5510	-0.9244
h1#01	0.5257	2.2279	-0.0603	2.8098
h1#02	1.5923	-1.8808	-1.2885	-2.7274
h1#03	-1.1333	1.0384	-0.4996	1.1762

TABLE V. Statistic data for ANN prediction results.

	ANN model for predicting intelligibility scores				ANN model for predicting intelligibility threshold				ANN model for predicting cadence threshold				ANN model for predicting audibility threshold			
	Tr ^a	Ve ^b	Te ^c	En ^d	Tr ^a	Ve ^b	Te ^c	En ^d	Tr ^a	Ve ^b	Te ^c	En ^d	Tr ^a	Ve ^b	Te ^c	En ^d
rms error	0.205	0.208	0.208	0.206	0.116	0.098	0.104	0.111	0.079	0.056	0.072	0.076	0.096	0.089	0.033	0.091
Error S.D. ^e	0.205	0.208	0.208	0.206	0.117	0.099	0.105	0.111	0.079	0.049	0.067	0.077	0.096	0.088	0.034	0.091
S.D. ^f ratio	0.477	0.476	0.475	0.476	0.280	0.246	0.244	0.267	0.192	0.115	0.139	0.183	0.254	0.211	0.096	0.240
R ²	0.773	0.773	0.774	0.773	0.921	0.940	0.941	0.929	0.963	0.987	0.982	0.967	0.936	0.958	0.992	0.942

^aTr=training set.

^bVe=verification set.

^cTe=test set.

^dEn=entire set.

^eError S.D.=standard deviation of the prediction error.

^fS.D. ratio=standard deviation ratio.

that of the original output data (an S.D. ratio of 0.1 is considered an excellent fit of the data, and an S.D. ratio of 1.0 is a poor fit).⁹ The correlation coefficients R^2 (the “coefficient of determination”) of the ANN predictions over the entire data sets for the intelligibility scores, intelligibility threshold, cadence threshold, and audibility threshold were 0.773, 0.929, 0.967, and 0.942, respectively. For comparison, in the work of Gover and Bradley,¹ the SII-weighted S/N measure has an R^2 of 0.762 for predicting the intelligibility scores and an R^2 of 0.919 for predicting the intelligibility threshold, and the S/N loudness ratio has an R^2 of 0.956 for predicting the cadence threshold and an R^2 of 0.899 for predicting the audibility threshold.

VII. DISCUSSIONS AND CONCLUSIONS

The present work indicates that the ANN approach provides a direct and accurate method for predicting speech intelligibility scores and security thresholds. The current method for predicting the speech intelligibility and privacy is first to develop a certain index and then relate the index to the subjective scores using a transfer function.^{1–3} The ANN approach can use the S/N ratio information to directly predict the subjective speech intelligibility score and security thresholds. Compared with the previous work¹ that used one-third-octave-band S/N ratios, the ANN approach produced comparable, or better in terms of the audibility threshold, prediction results using only the “octave band” S/N ratios.

Specifics of the ANN models for predicting the speech security and intelligibility scores are also provided in the present work. With this information, ANN models can be embedded into standard spreadsheet applications, thus allowing predictions to be made in a transparent and direct fashion.

Similar investigations that use one-third-octave-band (160 Hz–8 kHz) S/N ratios have also been conducted in the present work. However, the results of the one-third-octave-band analysis were only slightly better than those of the octave-band analysis. This may be due to the information in adjacent one-third-octave bands being highly correlated and therefore not contributing any significant new information to the ANN.

The present work only takes into account the S/N ratios in predicting the architectural speech intelligibility score and security thresholds. Other factors, such as talker gender and voice characteristics, speech material, room characteristics, and so on could be included in the ANN models to investigate the possibility of further improving the prediction accuracy.

ACKNOWLEDGMENTS

The authors are grateful to Professor Fergus R. Fricke at the School of Architecture, Design Science and Planning, University of Sydney, Australia, and Dr. Joseph Nannariello at Renzo Tonin & Associates Pty Ltd, Australia for their valuable comments on the present work. The present research was conducted under an Australian Postgraduate Award and a School of Architecture Design Science and Planning Supplementary Scholarship.

¹B. N. Gover and J. S. Bradley, “Measures for assessing architectural speech security (privacy) of closed offices and meeting rooms,” *J. Acoust. Soc. Am.* **116**, 3480–3490 (2004).

²ANSI S3.5-1969, “American National Standard Methods for the Calculation of the Articulation Index” (American National Standards Institute, New York).

³ANSI S3.5-1997, “American National Standard Methods for Calculation of the Speech Intelligibility Index” (American National Standards Institute, New York).

⁴J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation* (Addison-Wesley, Redwood City, CA, 1991).

⁵A. Browne, *Neural Network Analysis, Architectures and Applications* (Institute of Physics, Philadelphia, 1997).

⁶J. Nannariello, M. R. Osman, and F. R. Fricke, “Recent developments in the application of neural network analysis to architectural and building acoustics,” *Acoust. Aust.* **29**, 103–110 (2001).

⁷J. Nannariello and F. R. Fricke, “Introduction to neural network analysis and its application to building services engineering,” *Build. Services Eng. Res. Technol.* **22**, 58–68 (2001).

⁸F. F. Li and T. J. Cox, “Speech transmission index from running speech: A neural network approach,” *J. Acoust. Soc. Am.* **113**, 1999–2008 (2003).

⁹StatSoft, *STATISTICA Neural Networks version 4.0A manual* (StatSoft Inc, Tulsa, OK, 1999).

¹⁰J. Nannariello and F. R. Fricke, “A neural-computation method of predicting the early interaural cross-correlation coefficient (IACC_{E3}) for auditoria,” *Appl. Acoust.* **64**, 627–641 (2002).

¹¹J. Xu, J. Nannariello, and F. R. Fricke, “Predicting and optimising the airborne sound transmission of floor–ceiling constructions using computational intelligence,” *Appl. Acoust.* **65**, 693–704 (2004).

Estimates of auditory filter phase response at and below characteristic frequency (L)

Andrew J. Oxenham^{a)}

Research Laboratory of Electronics, and Speech and Hearing Bioscience and Technology Program, Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Stephan D. Ewert^{b)}

AG Medizinische Physik, Carl-von-Ossietzky Universität Oldenburg, 26111 Oldenburg, Germany

(Received 2 August 2004; revised 2 January 2005; accepted 7 January 2005)

Animal studies in basal cochlear regions have shown that basilar-membrane phase curvature (or rate of change of group delay with frequency) is negative around characteristic frequency (CF), but near zero well below CF. This study examined whether psychophysical masking experiments in humans show the same difference between on- and off-CF phase curvature. Masked thresholds were measured for a 2-kHz signal in the presence of harmonic tone complex maskers with a fundamental frequency of 100 Hz, band-limited between 200 and 1400 Hz (off-frequency masker) or between 1400 and 2600 Hz (on-frequency masker). The results from four normal-hearing listeners are consistent with predictions from animal physiological data: negative phase curvature is found for the on-frequency masker, whereas the phase curvature for the off-frequency masker is near zero. The method and results provide a strong test for the temporal response of computational models of human cochlear filtering. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1863012]

PACS numbers: 43.66.Ba, 43.66.Dc [JHG]

Pages: 1713–1716

I. INTRODUCTION

The amount of masking produced by a harmonic tone complex depends strongly on the phase relationships between the individual masker components (Smith *et al.*, 1986; Kohlrausch and Sander, 1995). The most commonly used stimuli in such experiments have been Schroeder-phase complexes (Schroeder, 1970), which have a group delay that changes linearly with frequency, producing a constant phase curvature and a time waveform that can be described as a repeating linear frequency glide. A modification of the original Schroeder equation, provided by Lentz and Leek (2001), which describes the phase of each component, is given by

$$\theta_n = C \pi n(n-1)/N, \quad (1)$$

where n is the component number, N is the total number of components, and C is a multiplicative constant. Positive (m_+) and negative (m_-) Schroeder-phase stimuli correspond to $C=1$ and $C=-1$, respectively, and a sine-phase (or cosine-phase) complex can be generated using $C=0$. Positive and negative Schroeder-phase complexes have very similar time waveforms,¹ but can produce differences in masked threshold that exceed 20 dB. The current explanation is that the positive phase curvature of the m_+ waveform interacts with the negative curvature of the basilar-membrane filter to produce a waveform with near-zero phase curvature, such as a sine-phase complex. The resulting highly modulated temporal envelope allows for signal detection at points in time with low signal-to-masker ratios. In contrast, the

negative phase curvature of the m_- waveform interacts with the basilar membrane to produce a waveform that still has a relatively flat temporal envelope, resulting in higher signal-to-masker ratios at threshold [Kohlrausch and Sander (1995); see also Oxenham and Dau (2001a), for the possible role of peripheral compression in determining thresholds in simultaneous masking]. The phase curvature of the complex is given by

$$\frac{d^2\theta}{df^2} = C \frac{2\pi}{Nf_0^2}, \quad (2)$$

where N is the total number of components in the complex, and f_0 is the fundamental frequency (Kohlrausch and Sander, 1995).

Using these complexes, it has been possible to derive estimates of the phase curvature of cochlear filtering at characteristic frequencies (CFs) between 125 and 8000 Hz (Oxenham and Dau, 2001b). The reasoning is that the masker phase curvature producing the lowest signal threshold is the one that produces the most highly modulated temporal envelope after cochlear filtering (Kohlrausch and Sander, 1995), which tends to occur when the filtered waveform has a phase curvature of zero. Thus, minimum threshold will be reached when the phase curvature of the masker is equal and opposite to the phase curvature of the cochlear filter at the CF corresponding to the signal frequency (Oxenham and Dau, 2001b). Physiological studies of basilar-membrane responses to Schroeder-phase stimuli have confirmed the basic pattern of expected effects, such as more modulated responses for stimuli with positive phase curvature (Recio and Rhode, 2000; Summers *et al.*, 2003).

Studies so far have used harmonic complexes with constant phase curvature (Smith *et al.*, 1986; Kohlrausch and

^{a)}Electronic mail: oxenham@mit.edu

^{b)}Present address: Center for Applied Hearing Research, Acoustic Technology, Ørsted-DTU, Technical University of Denmark, DK-2800 Lyngby, Denmark.

Sander, 1995; Carlyon and Datta, 1997; Summers and Leek, 1998; Oxenham and Dau, 2001a, b). By deriving the phase curvature of cochlear filtering from such maskers it is assumed that cochlear filtering can also be approximated as having constant phase curvature. As discussed by Oxenham and Dau (2001b), this assumption is supported by physiological data for frequencies around CF (Shera, 2001). However, at frequencies well below CF, phase curvature tends to zero. In fact the physiological data from animals using CFs of 1 kHz and above suggest that the phase response can be roughly divided into two regions, with negative curvature around CF and zero curvature for frequencies half an octave or more below CF (Shera, 2001). The situation appears to be different in apical regions of the cochlea (at CFs lower than about 1000 Hz), where the phase curvature at CF tends to zero (Oxenham and Dau, 2001b), or possibly even becomes positive (Carney *et al.*, 1999).

The change in phase curvature with frequency observed at high CFs provides an opportunity to further test the hypothesis that changes in masked threshold as a function of masker phase curvature reflect the phase response of cochlear filtering. Specifically, a masker with components around the signal frequency should produce minimum masking when the masker phase curvature is positive and “cancels out” the negative phase curvature of the on-frequency cochlear phase response. In contrast, a masker with components only well below the signal frequency should produce minimum masking when the masker phase curvature is zero, reflecting the fact that the off-frequency cochlear phase curvature is thought to be near zero. This prediction forms the basis of the experiment reported here.

II. METHODS

A. Stimuli

Masked thresholds of a 2-kHz sinusoid were measured in the presence of a harmonic tone complex masker with a fundamental frequency (F_0) of 100 Hz. The masker bandwidth extended from 1400 to 2600 Hz for the on-frequency condition, and from 200 to 1400 Hz for the off-frequency condition. Thus, the masker had the same bandwidth (1200 Hz) and the same total number of sinusoidal components (13) in both conditions. In the on-frequency condition, all masker components had equal amplitudes and the overall (rms) masker level was 73 dB SPL. In the off-frequency condition, the masker spectrum was shaped with a -6 dB/oct slope, such that each component was attenuated relative to the adjacent lower component, to reduce the influence of possible edge effects produced by the highest-frequency masker component (Kohlrausch and Houtsma, 1992). The overall (rms) off-frequency masker level of 88 dB SPL was 15 dB higher than in the on-frequency condition, so as to produce similar signal thresholds in both conditions. The signal had a total duration of 260 ms, including 30-ms raised-cosine onset and offset ramps, and was temporally centered within the masker, which had a total duration of 320 ms, including 10-ms raised-cosine onset and offset ramps. The starting phase of the signal was randomized from trial to trial. The starting phases of the masker components

were selected according to the modification of Schroeder's (1970) equation, as shown in Eq. (1) (Lentz and Leek, 2001), with C values ranging from -1 to 1 in steps of 0.25 .

The stimuli (masker and signal) were presented to the left ear of each subject. A contralateral Gaussian noise, band-pass filtered between 1400 and 2800 Hz, was gated synchronously with each masker and presented to the right ear at a spectrum level of 20 dB SPL/Hz. This was to prevent the detection of the signal by the right ear via acoustic or electric crosstalk. In the on-frequency conditions, a pink noise, band-pass filtered between 25 and 1000 Hz, was presented to the left (signal) ear at a level of 44 dB SPL per $\frac{1}{3}$ octave band. This level was chosen to be 25 dB below the level of the masker within the $\frac{1}{3}$ octave around the signal frequency. The purpose of the pink noise was to mask possible distortion products generated by the signal and masker.

Stimuli were generated digitally and played out via a RME DIGI96/8 PAD soundcard and an external SEKD ADSP 2496 PRO digital-to-analog converter at 24-bit resolution and a sampling rate of 32 kHz. The stimuli were then passed through a Behringer HA4600 headphone buffer and presented over Sennheiser HD 580 headphones to listeners seated in a double-walled sound-attenuating booth.

B. Procedure

An adaptive three-interval three-alternative forced-choice procedure was used in conjunction with a two-down one-up tracking rule to estimate the 70.7% correct point on the psychometric function (Levitt, 1971). The masker was presented on each interval and the signal was presented randomly in one of the three intervals with equal probability. Each interval in a trial was separated by an interstimulus interval (ISI) of 500 ms. The intervals were marked on a computer monitor and feedback was provided after each trial. Listeners responded via the computer keyboard or mouse. The initial step size was 8 dB, which was reduced to 4 dB after the first two reversals and then to 2 dB after the next two reversals. Threshold was defined as the mean of the remaining six reversals. At least four threshold estimates were made for each condition. In the very rare case that standard deviation across the four runs was greater than 4 dB, another two estimates were made and the mean of all six was recorded. The conditions were run using a randomized blocked design, with all conditions being presented once before embarking on a repetition of the conditions. The order of presentation of the conditions was selected randomly for each listener and each repetition. Measurements were made in 2-h sessions, including many short breaks. No more than one session per listener was completed in any one day.

C. Listeners

Four listeners (three male, one female—ML) participated. Two were the authors; the other two were students at the University of Oldenburg. The ages of the subjects ranged from 23 to 32 years. All had audiometric thresholds of 15 dB HL or lower at octave frequencies between 250 and 8000 Hz. Subjects AO, SE, and KB had considerable experience in psychoacoustic detection tasks and were familiarized with

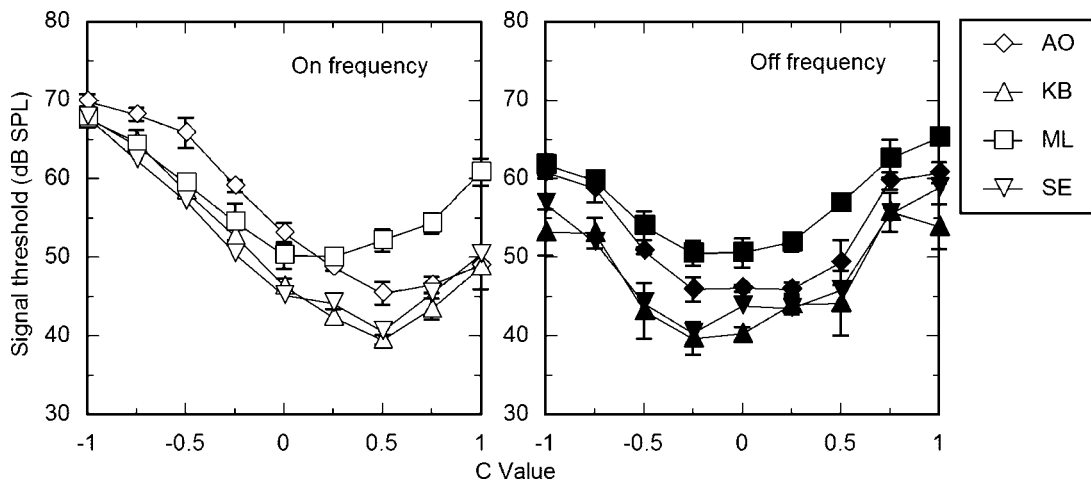


FIG. 1. Individual signal thresholds in the presence of the on-frequency (left panel) and off-frequency (right panel) maskers. Error bars denote ± 1 standard deviation across the individual repetitions. In line with predictions, masking minima are reached for a positive phase curvature for on-frequency masker, but are closer to zero for the off-frequency masker.

the present experiment before data were collected. Subject ML had no previous experience in psychoacoustic experiments and received about 2 h of training before data collection started. Subjects KB and ML were compensated for their services on an hourly basis.

III. RESULTS AND DISCUSSION

The individual results are shown in Fig. 1. The left panel shows results with the on-frequency masker; the right panel shows results with the off-frequency masker. Different symbols represent thresholds from different subjects, as shown in the legend, and error bars denote ± 1 standard deviation of the individual estimates. Despite differences in individual thresholds of 10 dB or more, the patterns of results are generally similar across the four subjects. In particular, all four subjects show a minimum masked threshold at a positive C value for the on-frequency masker: AO, KB, and SE show a minimum at around $C=0.5$, while ML shows similar thresholds at $C=0$ and $C=0.25$, with a slightly lower value at 0.25. All four subjects show a minimum at a lower C value for the off-frequency masker: individual minima are either at $C=0$ or $C=-0.25$.

The differences between the on- and off-frequency conditions are seen clearly in a direct comparison of the mean results (Fig. 2). The two curves are fitted sinusoidal functions, as described in Oxenham and Dau (2001b). The minima of the fitted functions were 0.42 and -0.05 for the on- and off-frequency conditions, respectively. The masker curvature at these minima in rad/Hz^2 [as calculated from Eq. (2)] can be normalized into dimensionless units by multiplying by $f_s^2/2\pi$, where f_s is the signal frequency (Oxenham and Dau, 2001b; Shera, 2001). The normalized masker curvature at the minimum of the masking function for the present on-frequency data is 12.9, implying a normalized cochlear filter phase curvature of -12.9 , which is in good agreement with previous estimates at 2000 Hz, using different subjects and different masker F0s and bandwidths, which range between -8 and -16 (Lentz and Leek, 2001; Oxenham and Dau, 2001b). The novel condition tested here involves the off-frequency masker. Here, the estimated curva-

ture using the same signal frequency is close to zero, in clear contradiction to an assumption that curvature remains constant with frequency at a given CF.

Overall, the data support the hypothesis outlined in the introduction:

- (1) Masking patterns for the off-frequency masker produce a minimum at around zero phase curvature, suggesting no phase curvature in the cochlear filter response to stimuli well below CF.
- (2) Masking patterns for the on-frequency masker produce a minimum at a positive masker phase curvature, presumably counteracting the negative phase curvature of cochlear filtering around CF.

A phase versus frequency plot of these data (not shown) is consistent with physiological data of basilar-membrane motion at the basal end of the cochlear, in showing a nearly straight-line section for frequencies half-an-octave or more below CF and a more curved negative-going section for frequencies around CF (de Boer and Nuttall, 1997; Ruggero *et al.*, 1997; Rhode and Recio, 2000; Shera, 2001). The situ-

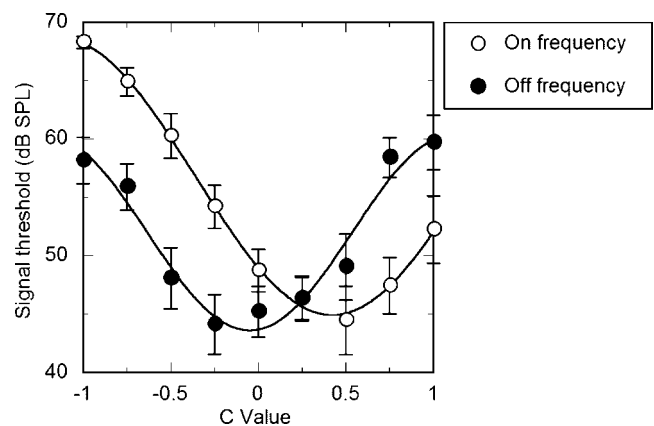


FIG. 2. Mean signal thresholds, along with fitted sinusoidal functions. Open symbols represent thresholds from the on-frequency condition; filled symbols represent thresholds from the off-frequency condition. Error bars denote ± 1 standard error of the mean across the four subjects. Minima of the fitted functions are reached for C values of 0.42 and -0.05 for the on- and off-frequency conditions, respectively.

ation at signal frequencies well below the one tested here (2000 Hz) is likely to be different, in that the phase curvature tends to zero even in the on-frequency case (Oxenham and Dau, 2001b).

One assumption of our technique is that the masker phase curvature that produces the most modulated temporal envelope is independent of the relative component amplitudes. This assumption is important because component amplitudes can be varied by experimental manipulations as well as by outer-, middle-, and inner-ear filtering. Numerical simulations have confirmed this assumption in all conditions tested so far when the Hilbert envelope is used. However, the assumption does not always hold when measures such as the crest factor (Strickland and Viemeister, 1996) are used directly on the stimulus waveform, rather than the envelope. Although it is universally assumed that signal detection in these situations is based on envelope, rather than raw waveform, properties, we cannot completely rule out some influence of temporal factors not represented in the Hilbert envelope. Empirical tests are required to settle this question, although informal testing has not yet yielded any situations where the phase curvature producing minimum threshold is affected by the relative amplitudes of the masker components.

One potential limitation of the present technique is that it does not allow a detailed view of how the curvature changes with frequency; we have only shown the estimated curvature for two fairly broad sample frequency regions. By using maskers with nonconstant phase curvature, it may be possible to map out changes in curvature with frequency in more detail. However, it may be that inherent measurement errors would swamp more fine-grained effects than those found here. A method for using envelope modulation in the auditory nerve to map out cochlear phase responses has been proposed recently (van der Heijden and Joris, 2003). By using the modulation at the beat frequencies between unequally spaced sinusoids, it was possible to derive a much more detailed phase response even in cochlear regions where the auditory nerve fibers no longer phase-lock to the temporal fine structure. While our technique and that of van der Heijden and Joris (2003) both use aspects of envelope modulation, theirs has the advantage of providing a much more detailed function of phase versus frequency. Unfortunately, there does not seem to be any obvious way of translating their technique into a task suitable for behavioral experiments.

In summary, masked thresholds for a sinusoidal signal in a harmonic tone complex vary with masker phase curvature and masker frequency region as would be predicted from animal physiological studies of basal cochlear mechanics: on-frequency maskers indicate negative phase curvature, while maskers well below the signal frequency indicate near-zero phase curvature. The results suggest that it is possible—to a limited extent—to map out changes in human cochlear phase response with frequency. Such results should provide important tests for computational models of human cochlear filtering.

ACKNOWLEDGMENTS

This work was supported primarily by the National Institutes of Health (Grant No. R01 DC 03909). It was begun while the second author was a visiting scientist at MIT's Research Laboratory of Electronics, and was completed while the first author was a fellow at the Hanse Institute for Advanced Study, Delmenhorst, Germany. We thank Torsten Dau, Van Summers, and an anonymous reviewer for helpful comments on previous versions of this manuscript.

¹If cosines are used as the basis functions, the negative sign in the equation leads to a simple time reversal of the overall waveform. If sines are used as the basis functions (as was done in this study), the negative sign leads to a time-reversal and inverting of the overall waveform. This is because the sine and cosine functions are dot-symmetric and symmetric about zero, respectively.

- Carlyon, R. P., and Datta, A. J. (1997). "Masking period patterns of Schroeder-phase complexes: Effects of level, number of components, and phase of flanking components," *J. Acoust. Soc. Am.* **101**, 3648–3657.
- Carney, L. H., McDuffy, M. J., and Shekter, I. (1999). "Frequency glides in the impulse responses of auditory-nerve fibers," *J. Acoust. Soc. Am.* **105**, 2384–2391.
- de Boer, E., and Nuttall, A. L. (1997). "The mechanical waveform of the basilar membrane. I. Frequency modulations (glides) in impulse responses and cross-correlation functions," *J. Acoust. Soc. Am.* **101**, 3583–3592.
- Kohler, A., and Houtsma, A. J. M. (1992). "Pitch related to spectral edges of broadband signals," *Philos. Trans. R. Soc. London, Ser. B* **336**, 375–382.
- Kohler, A., and Sander, A. (1995). "Phase effects in masking related to dispersion in the inner ear. II. Masking period patterns of short targets," *J. Acoust. Soc. Am.* **97**, 1817–1829.
- Lentz, J. J., and Leek, M. R. (2001). "Psychophysical estimates of cochlear phase response: Masking by harmonic complexes," *J. Assoc. Res. Otolaryngol.* **2**, 408–422.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Oxenham, A. J., and Dau, T. (2001a). "Reconciling frequency selectivity and phase effects in masking," *J. Acoust. Soc. Am.* **110**, 1525–1538.
- Oxenham, A. J., and Dau, T. (2001b). "Towards a measure of auditory-filter phase response," *J. Acoust. Soc. Am.* **110**, 3169–3178.
- Recio, A., and Rhode, W. S. (2000). "Basilar membrane responses to broadband stimuli," *J. Acoust. Soc. Am.* **108**, 2281–2298.
- Rhode, W. S., and Recio, A. (2000). "Study of mechanical motions in the basal region of the chinchilla cochlea," *J. Acoust. Soc. Am.* **107**, 3317–3332.
- Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (1997). "Basilar-membrane responses to tones at the base of the chinchilla cochlea," *J. Acoust. Soc. Am.* **101**, 2151–2163.
- Schroeder, M. R. (1970). "Synthesis of low peak-factor signals and binary sequences with low autocorrelation," *IEEE Trans. Inf. Theory* **16**, 85–89.
- Shera, C. A. (2001). "Frequency glides in click responses of the basilar membrane and auditory nerve: Their scaling behavior and origin in traveling-wave dispersion," *J. Acoust. Soc. Am.* **109**, 2023–2034.
- Smith, B. K., Sieben, U. K., Kohler, A., and Schroeder, M. R. (1986). "Phase effects in masking related to dispersion in the inner ear," *J. Acoust. Soc. Am.* **80**, 1631–1637.
- Strickland, E. A., and Viemeister, N. F. (1996). "Cues for discrimination of envelopes," *J. Acoust. Soc. Am.* **99**, 3638–3646.
- Summers, V., and Leek, M. R. (1998). "Masking of tones and speech by Schroeder-phase harmonic complexes in normally hearing and hearing-impaired listeners," *Hear. Res.* **118**, 139–150.
- Summers, V., de Boer, E., and Nuttall, A. L. (2003). "Basilar-membrane responses to multicomponent (Schroeder-phase) signals: Understanding intensity effects," *J. Acoust. Soc. Am.* **114**, 294–306.
- van der Heijden, M., and Joris, P. X. (2003). "Cochlear phase and amplitude retrieved from the auditory nerve at arbitrary frequencies," *J. Neurosci.* **23**, 9194–9198.

Design and optimization of a noise reduction system for infrasonic measurements using elements with low acoustic impedance

Benoit Alcoverro^{a)} and Alexis Le Pichon^{b)}

CEA/DASE-LDG, BP12, 91680 Bruyères-le-Châtel, France

(Received 22 July 2002; revised 17 March 2004; accepted 29 March 2004)

The implementation of the infrasound network of the International Monitoring System (IMS) for the enforcement of the Comprehensive Nuclear-Test-Ban Treaty (CTBT) increases the effort in the design of suitable noise reducer systems. In this paper we present a new design consisting of low impedance elements. The dimensioning and the optimization of this discrete mechanical system are based on numerical simulations, including a complete electroacoustical modeling and a realistic wind-noise model. The frequency response and the noise reduction obtained for a given wind speed are compared to statistical noise measurements in the [0.02–4] Hz frequency band. The effects of the constructive parameters—the length of the pipes, inner diameters, summing volume, and number of air inlets—are investigated through a parametric study. The studied system consists of 32 air inlets distributed along an overall diameter of 16 m. Its frequency response is flat up to 4 Hz. For a 2 m/s wind speed, the maximal noise reduction obtained is 15 dB between 0.5 and 4 Hz. At lower frequencies, the noise reduction is improved by the use of a system of larger diameter. The main drawback is the high-frequency limitation introduced by acoustical resonances inside the pipes.

© 2005 Acoustical Society of America. [DOI: 10.1121/1.1804966]

PACS numbers: 43.20.Dk, 43.20.El, 43.28.Bj, 43.28.Dm [LCS]

Pages: 1717–1727

I. INTRODUCTION

A renewed interest in the detection of infrasonic energy in the atmosphere appeared with the CTBT Organization. Associated with seismic, hydroacoustic, and radionuclide monitoring, the infrasound detection systems are dedicated to the monitoring of atmospheric tests. Ever since their progressive installation, these systems have also shown their importance in the monitoring of the environment, through the study of low-frequency acoustic signals of both natural and human origin.

The main problem in the detection of infrasonic waves is the reduction of acoustic background noise related to atmospheric turbulence. One way of attenuating the noise consists of connecting the pressure sensor to spatial acoustic filters. These filters make use of the incoherent structure of the background noise compared with the infrasonic pressure wave (McDonald and Herrin, 1975; Bedard, 1977). The turbulent movement of air close to the ground is complex. In a general way, atmospheric turbulence at a point consists of eddies moved by the wind that set the mass of air in rotation. The largest diameter (D) of the eddy is linked to the frequency of the pressure fluctuations (f) and the wind speed (V) by the well-know relation $f=V/D$. In the case of coherent acoustic waves, the ratio between the propagation velocity of the wind-related noise and the propagation speed of the infrasonic signal is much lower than unity. For a given frequency, this implies that several wavelengths of atmospheric noise exist in the same space occupied by only one wavelength of the acoustic signal. Assuming the wavelength of

the pressure to be measured is large compared with the dimensions of the filtering system, the summation of pressure—comprising both acoustic signal and atmospheric noise—from a spatial array of air inlets leads to a noticeable improvement in the signal-to-noise ratio. In the case of random noise with normal distribution amplitude, the differences of phase between the various points of measurement reduce the amplitude of the noise by a factor equal to the square root of the number of measurement points (Daniels, 1959). In reality, the performance of the filtering system is limited by attenuation and distortion during propagation of the signal in the pipes. In particular, the natural resonance of the system becomes a major disadvantage with increasing dimension of the system.

The majority of existing systems are based on pipes proposed by Daniels (Daniels, 1959). This system is made up of a succession of rigid pipes of increasing diameter, forming a linear array, to which is connected capillary air inlets. A satisfactory signal-to-noise ratio is attained for frequencies lower than 1 Hz. However, linear systems (i.e., with constant attenuation of sound propagation per unit length) exhibit a nonisotropic noise reduction and sensitivity (Noel and Whitaker, 1991). A circular system with a more omnidirectional response is better adapted, such as that proposed by Burridge (1971) and Grover (1971).

A qualitative analysis of measurements from different systems underlines the fact that long linear systems can be replaced by circular systems of a smaller size (Noel and Whitaker, 1991; Grover, 1971). With a diameter of 100 m and for frequencies lower than 0.5 Hz, these systems display performances comparable with the linear system of Daniels. Circular systems have the advantage of showing isotropic noise reduction when the impedance of the air inlets is suf-

^{a)}Electronic mail: benoit.alcoverro@antigone.cea.fr

^{b)}Electronic mail: alexis.le-pichon@cea.fr

ficiently large to allow perfect acoustic summation over the entire circumference of the circle (Burrige, 1971). However, in addition to the problems of maintenance related to the large size of these systems, with particular care to avoid obstructing the air inlets, a major disadvantage is the difference in response between each air inlet for frequencies of the order of 1 Hz as well as the possible resonances induced in the [0.02–4] Hz band.

Other noise reduction systems are based on the use of microporous pipes arranged radially in a circle centered on the microbarometer (Noel and Whitaker, 1991). The performances of a 16-m diameter ring fitted with 96 capillary air inlets and a system with porous pipes arranged in a 25-m cross are comparable. The porous pipe solution is inexpensive and provides a significant improvement of the signal-to-noise ratio in the [0.1–10] Hz frequency band (Haak and de Wilde, 1996). However, the performances of these systems deteriorate rather quickly with time under severe meteorological conditions. The major drawback of the porous pipes and the capillaries are the variability of the physical properties with time under standard weather conditions (prolonged exposure to bad weather and dust). Although this solution is attractive as well as efficient, it is not sufficiently robust to be used under the operational conditions of the CTBT detection networks.

Our aim in this paper is the dimensioning and optimization of a discrete acoustical-mechanical system using elements with low acoustic impedance by using numerical simulations (Alcoverro, 1998). This system presents roughly the same performances obtained with porous pipes. A precise electroacoustic modeling of the system takes into account the frequency response of the acoustical circuits and the sensor coupling effects. A synthetic input pressure is generated by a wind model based on the analysis of the spatial coherence of the background noise as a function of wind speed. These numerical simulations, validated by a limited number of experimental measurements, provide an optimum diameter and number of air inlets in the [0.02–4] Hz frequency band under different wind conditions.

II. ACOUSTIC MODELING OF A DISCRETE MECHANICAL SYSTEM

A. Description of the parameters and dimensioning

This system consists of nonporous pipes and low-impedance air inlets that are directly connected to the microbarometer. The air inlets are composed of open pipes protected by a fine mesh to avoid an obstruction. To maintain isotropic noise reduction, the distribution of the air inlets is made as circular as possible. Such geometry ensures an identical transfer function between each air inlet and the microbarometer (Fig. 1). In order to reduce the complexity of the system, each air inlet is connected to the sensor by two pipes of length L_1 and L_2 and also by means of four primary summation cavities. The measured and calculated improvement in the signal-to-noise ratio is of the order of 15 dB in the [0.1–10] Hz band (Alcoverro, 1998). The diameter of the system depends on the noise reduction expected at low frequency for a given wind speed, as well as on the upper limit

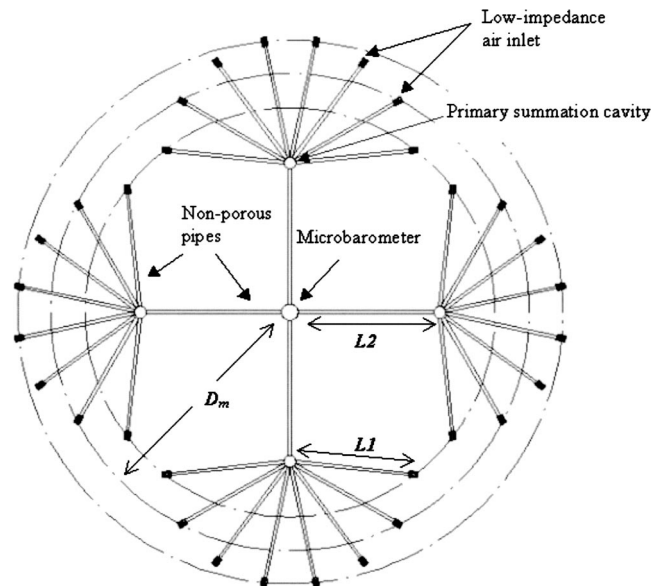


FIG. 1. Discrete mechanical noise reduction system, composed of 32 low-impedance air inlets, arranged in a circle of 16 m mean diameter (D_m) with $L_1=4$ m and $L_2=5$ m. The proposed geometry makes it possible to obtain an identical travel time between each air inlet and the sensor. The distance between the air inlets is always the same. Beyond 1 Hz, the maximum noise reduction is of the order of 15 dB for wind speeds lower than 2 m/s. The first resonance frequency is at 12 Hz.

imposed by the first resonance frequency of the pipes (Sec. IV). A diameter ranging between 10 and 20 m yields a first resonance frequency around 10 Hz, which ensures a flat response in amplitude up to 4 Hz.

For the simulation, we first construct an electroacoustic model from a given geometry (mean diameter and number of air inlets) and the constructive parameters (air entries, pipes, summation cavities, and the model of the microbarometer used). The transfer functions $H_i(\omega)$ linking each air inlet with the pressure measured by the sensor is calculated. The principle of superposition is used to calculate the total response of the system by temporal summation, considering the other air inlets as open. Then, the influences on the amplitude response of the length and diameters of the pipes, the cavity volumes, and the number of air inlets are determined. The noise reduction of a system subjected to a given wind speed is obtained by convolution of the pulse response $h_i(t)$ of all the transfer functions $H_i(\omega)$ with the synthetic wind-noise pressure at each air inlet $w_i(t)$ (Sec. III B). To incorporate the concept of spatial filtering into the model, $h_i(t)$ takes account of the propagation delay (from the air inlet i to a reference point) introduced by the propagation speed of eddies (controlled by the wind speed). The noise reduction is obtained by calculating the spectra ratio of $\sum_i \{h_i(t) * w_i(t)\}$ and the synthetic wind-noise $w_i(t)$ (Sec. IV A).

B. Electroacoustic modeling

The frequency response of the system in amplitude and phase is obtained by calculating the transfer functions $H_i(\omega)$. Figure 2 presents an acoustic analog model of the noise reducer shown in Fig. 1. The elements constituting this model are given in Appendix A, including the air inlet, the

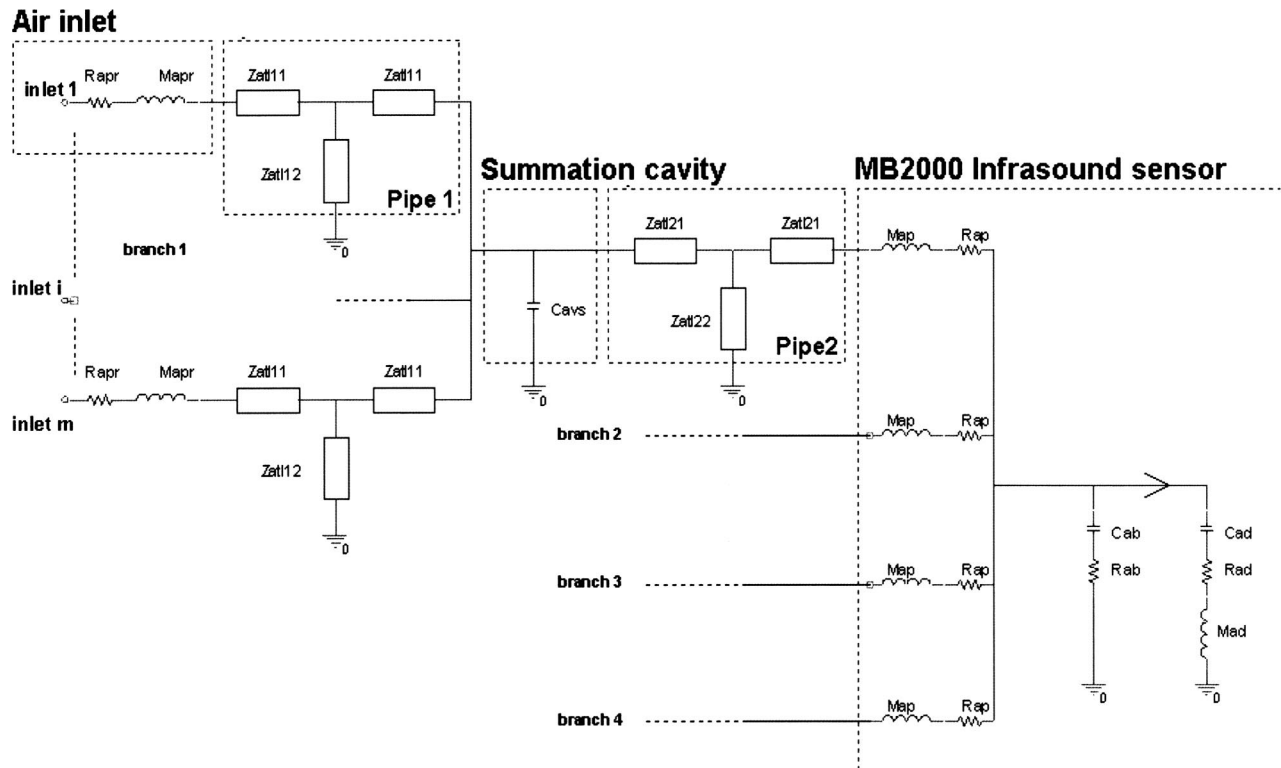


FIG. 2. Electroacoustic model of a noise reducer with 32 air inlets and 4 branches.

pipes, the summation cavity, and the microbarometer used. The method of calculation of the transfer functions is given in Appendix B.

C. Optimization of the constructive parameters

In this section, the frequency response of the noise reduction system is analyzed through a parametric study. Figure 3 summarizes the main effects induced by a change of the following parameters: average diameter $D_m = 2(L_1 + L_2)$ and internal diameter of the pipes $2r_p$, number of air inlets and volume of the summation cavities V_s (the 0 dB reference is the level of excitation provided by the signal applied to each air inlet).

The main results are as follows.

- (i) Figure 3(a) shows several resonances above 1 Hz. These resonances are mainly due to the length of the pipes terminated by a low impedance (pipes are open or emerge into a cavity). The frequency peaks corresponds to the fundamentals and odd harmonics resonances in each elementary pipe. If the impedance at the end of each pipe had a value close to the characteristic impedance of an infinite pipe, namely $\rho c/S$, these resonances could be reduced. With an internal diameter of 15 mm, the value of the characteristic impedance is $2.3 \times 10^6 \Omega \text{A}$ and could be obtained by using a 50-mm length capillary with 2 mm internal diameter. The frequency f_0 of the first resonance is inversely proportional to the average diameter D_m of the system and depends on the sound speed c . f_0 has

been evaluated as

$$f_0 \sim 0.7 \frac{c}{D_m}. \quad (1)$$

- (ii) As shown by Fig. 3(b), any increase in the internal diameter tends to decrease the viscothermic losses, which also amplifies the resonances. On the other hand, there is little change in the resonant frequency (because of the decreasing acoustic mass of the pipe). Using a small diameter increases the risk of obstruction by foreign bodies or condensation. The choice of an internal diameter around 15 mm seems reasonable. This value is very close to the diameter of the air entries on the sensor, allowing a correct acoustic impedance match at the entry to the sensor.
- (iii) Figure 3(c) shows that the frequency of the resonances is unaffected by the number of air inlets. Only the gain of the transfer function is reduced as a function of the number of inlets (about -33 dB for 48 air inlets). After summation of all the measurement points, a signal at 4 Hz is not attenuated (Sec. IV E).
- (iv) Figure 3(d) confirms that the volume of the summation cavities should not exceed 1 l. If this volume is higher, the equivalent capacity of the cavity increases, which tends to decrease the cutoff frequency of the low pass circuit formed by the acoustic mass of the pipes combined with this capacity.

Figure 4 presents the transfer function of the 16-m noise reducer. The amplitude response is uniform up to 5.5 Hz (± 3 dB), while the phase does not vary by more than 10° below

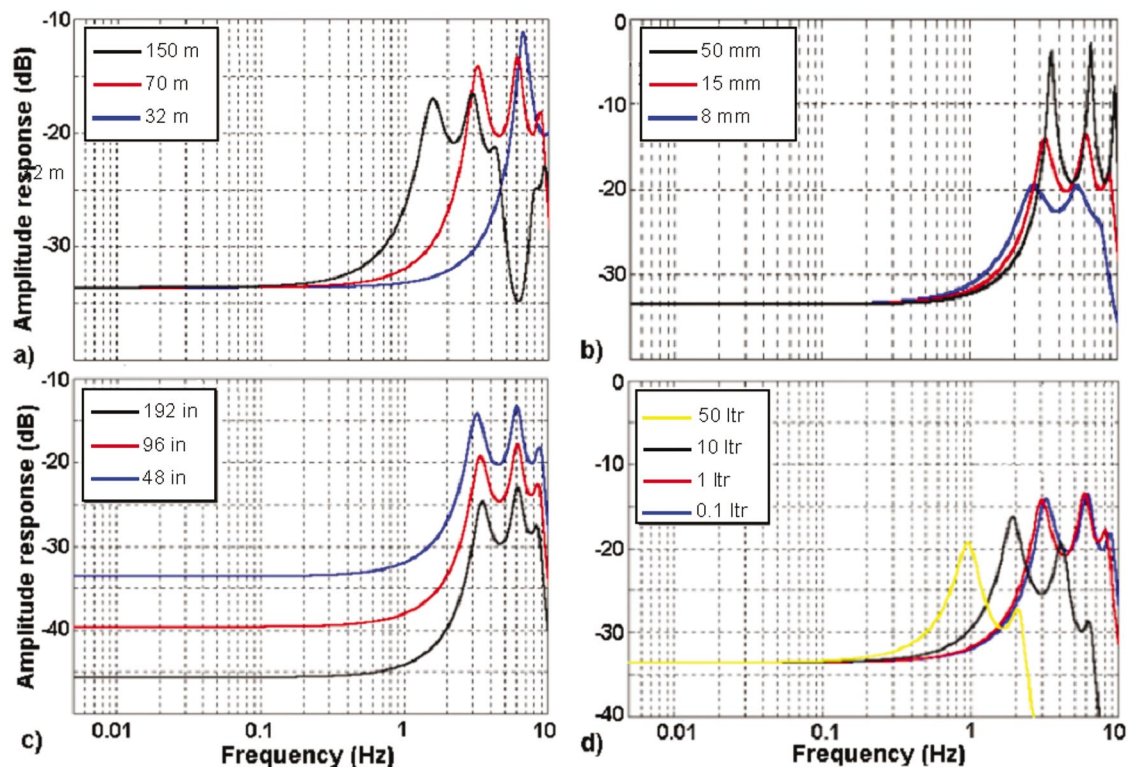


FIG. 3. Effects of the constructive parameters on frequency response. (a) Influence of the pipe length l : system with 48 air inlets and pipes of 15 mm internal diameter. (b) Influence of the internal diameter: 48 air inlets and 70 m length. (c) Influence of the number of air inlets: pipes of 15 mm diameter and 70 m length. (d) Influence of the volume of the summation cavities: 0.1 to 50 l for 48 air inlets and 70 m length.

4 Hz. The performance of such a system is limited to low wind speeds (Sec. IV E).

III. ESTABLISHMENT OF A WIND MODEL

The simulation of the noise reduction requires the establishment of a model representative of the wind-noise pres-

sure. This model is derived from observations of the coherence criterion for atmospheric pressure fluctuations measured at several points.

A. Measurements of coherence

It has been shown that, for periods ranging between and 0.2 s and 1 week, the atmospheric pressure spectrum is

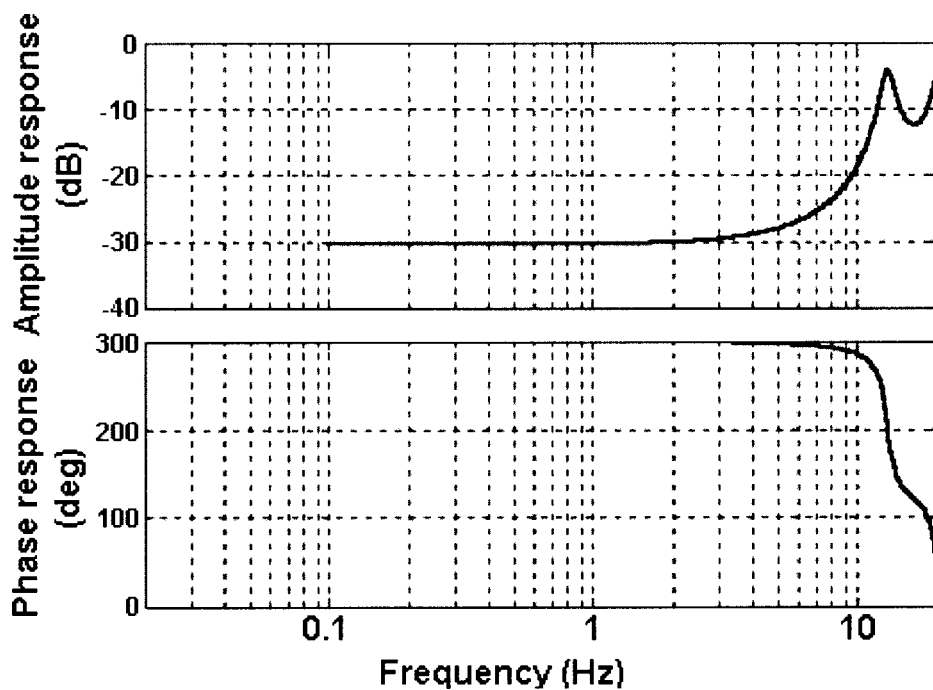


FIG. 4. Transfer function of the 16 m system.

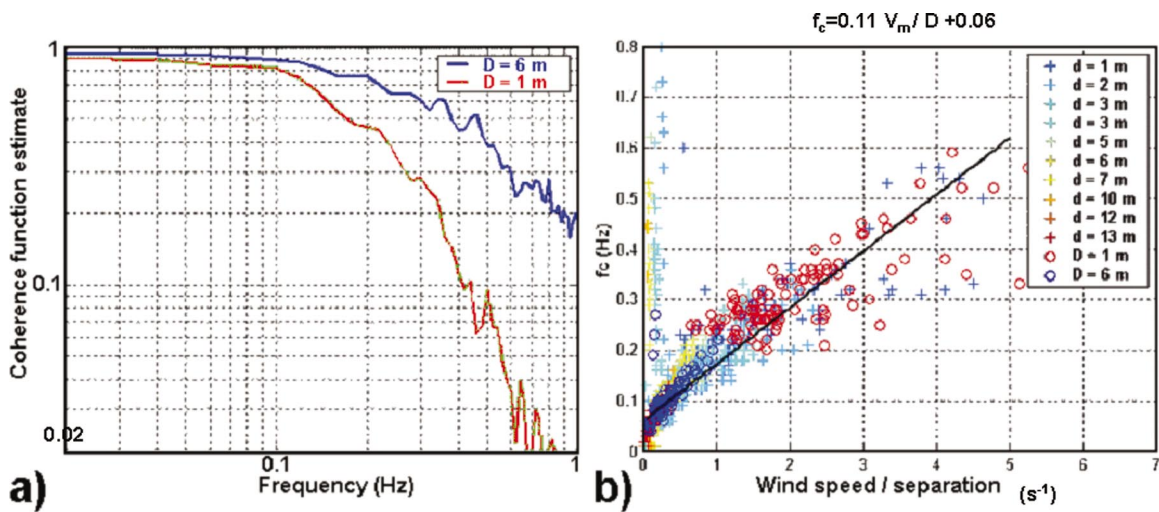


FIG. 5. Coherence as a function of frequency and wind speed. (a) Coherence plotted as a function of frequency for wind speed of 5 m/s and distances of 1 and 6 m. (b) Frequency of decorrelation plotted as a function of intersensor distance and wind speed.

strongly dependent on gravity waves as well as convective effects in the boundary layer of the atmosphere (Gossard and Munk, 1954; Canavero and Einaudi, 1987; Einaudi *et al.*, 1989). The amplitude of the wind-generated noise is strongly correlated with wind speed at the surface for periods shorter than 300 s, with most of the energy concentrated below 60 s (Kimball and Lemon, 1970; Grover, 1971; Wilczak *et al.*, 1992). In particular, hydrostatic and dynamic effects have been identified as the principal physical phenomena producing the pressure fluctuations (Bedard *et al.*, 1992). The pressure spectrum is proportional to a power of the frequency, with the exponent ranging between $-5/3$ and $-7/3$ (Eliot, 1972; Canavero and Einaudi, 1987; Wilczak *et al.*, 1992).

Experiments were carried out to refine the coverage of the parameters controlling the spatial correlation of the pressure variations at low wavelengths (1–10 m). The decrease of coherence with increasing distance between two microbarometers is more marked in the direction of the prevailing winds than in the transverse direction (McDonald *et al.*, 1971). In order to reduce the wind direction effect on the coherence measurements, the pressure variations were measured on cleared ground (covered with close-cropped grass) by five microbarometers set up in two configurations:

- (i) sensors aligned in the direction of the prevailing winds;
- (ii) sensors laid out in two overlapping equilateral triangles, with sides of 1 and 6 m, having one of the apices in common. The anemometer is fixed at a height of 1.5 m in the center of the larger triangle.

For eight days, the pressure, wind speed, and direction were continuously recorded using a 24 bits DASE digitizer at a 20 Hz sample rate.

The decorrelation of the background noise has been evaluated using the standard procedure for coherence estimation (Priestley, 1966; Mack and Flinn, 1971; McDonald and Herrin, 1975; Einaudi *et al.*, 1989). The dependence of the background noise coherence has been studied as a function of the distance separating the sensors for wind speeds rang-

ing from 0 to 7 m/s. The frequency above which the coherence is lower than 0.5, named decorrelation frequency f_c , has been extracted from measurement data on 500 s time windows. Figure 5(a) gives an example of the variation of coherence for two different distances. These measurements show that convected turbulence is normally dispersive. When measurements are taken by sensors separated by some fraction of the eddy diameters, coherence decreases with increasing distance between sensors. Figure 5(b) shows that, to a first approximation, f_c increases as a function of the V_m/D ratio, where V_m is the average wind speed and D is the distance separating the measurement points. A first-order curve fitting yields

$$f_c \approx 0.11 \frac{V_m}{D} + 0.06. \quad (2)$$

(In a future work, more exact fitting using polynomial fittings and a larger wind speed data recording will give more precise results). In addition to the wind speed, atmospheric turbulence also depends on the properties of the boundary layer, the nature of the surface hydrostatic effects, and the presence of atmospheric waves (Mack and Flinn, 1971; McDonald and Herrin, 1975; Einaudi *et al.*, 1989). The measured bias and the scattering of measurements may be ex-

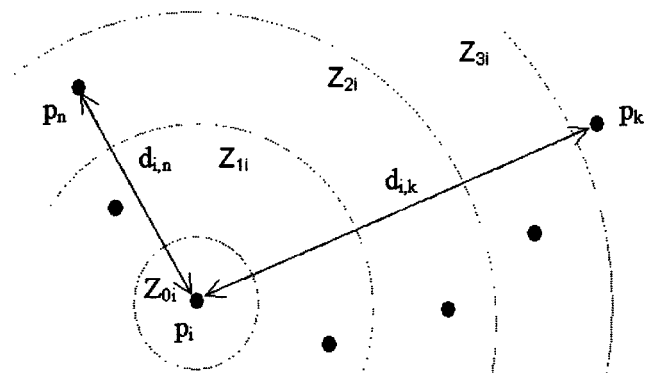


FIG. 6. Definition of the coherence area.

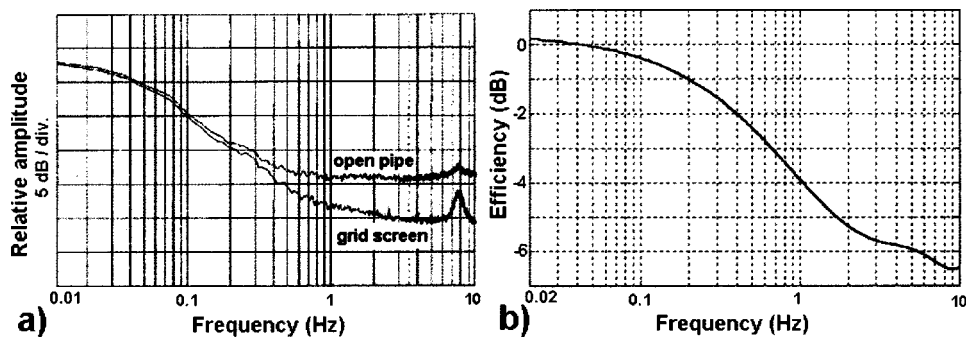


FIG. 7. Effect of grid screens on noise reduction. (a) Comparative spectra of the noise generated in a wind tunnel operated at 2 m/s, measured with and without grid screens (the rotational frequency of the blades is 8 Hz). (b) Curve of the $P(f)$ polynomial modeling the noise reduction obtained with grid screens.

plained by the complex structure of atmospheric turbulence phenomena. A greater dispersion of the points is noted for wind speeds lower than 1 m/s. This is attributed to a marked stable peak with a period between 4 and 7 s, originating from oceanic swell (Donn, 1967). The strong correlation of these signals from one sensor to another partly explains the increase of the decorrelation frequency between 0.2 and 0.5 Hz at low wind speeds.

B. Wind model

If n measurement points p_i are considered, $n-1$ points p_k are at a distance of $d_{i,k}$ from point p_i . Between points p_k and p_i , the wind-noise pressure exhibits a coherence function whose cut-off frequency is given by Eq. (2). Correlation area Z_{ni} attached to points p_i is shown in Fig. 6. These zones of coherence Z_{ni} are defined like circular areas centered on p_i with a radius equal to an entire multiple of the smallest $d_{i,k}$. Several points p_k can be present within the zones of coherence. To satisfy Eq. (2), a normal random temporal sequence $\{n_i\}$ is applied to point p_i . The distance $d_{i,k}$ is calculated for each point p_k and the sequence $\{n_{i,k}\}$ is applied to points p_k . The $\{n_{i,k}\}$ sequences are derived from the $\{n_i\}$ sequences by using a first-order low-pass filter with a cut-off frequency of

$$fc_{i,k} = 0.2 \frac{V_m}{d_{i,k}}. \quad (3)$$

For each point p_i , the signal representative of the wind is generated by

$$\{w_i(t)\} = \sum_{k=1 \dots n} \{n_{i,k}(t)\}. \quad (4)$$

The procedure is repeated for all points p_i with $i=[1,n]$. The spectrum of the wind-noise pressure thus generated presents a slope close to -1.3 , which is reasonably close to the experimental values.

IV. SIMULATION OF NOISE REDUCTION AND EXPERIMENTS

The impulse responses $h_i(t)$ are derived from the transfer function $H_i(\omega)$ established in Sec. II. These responses $h_i(t)$ are convolved with the synthetic wind noise $w_i(t)$. The response of the system to a wind excitation is given by the sum of the convolution:

$$y(t) = \sum_i \{h_i(t) * w_i(t)\}. \quad (5)$$

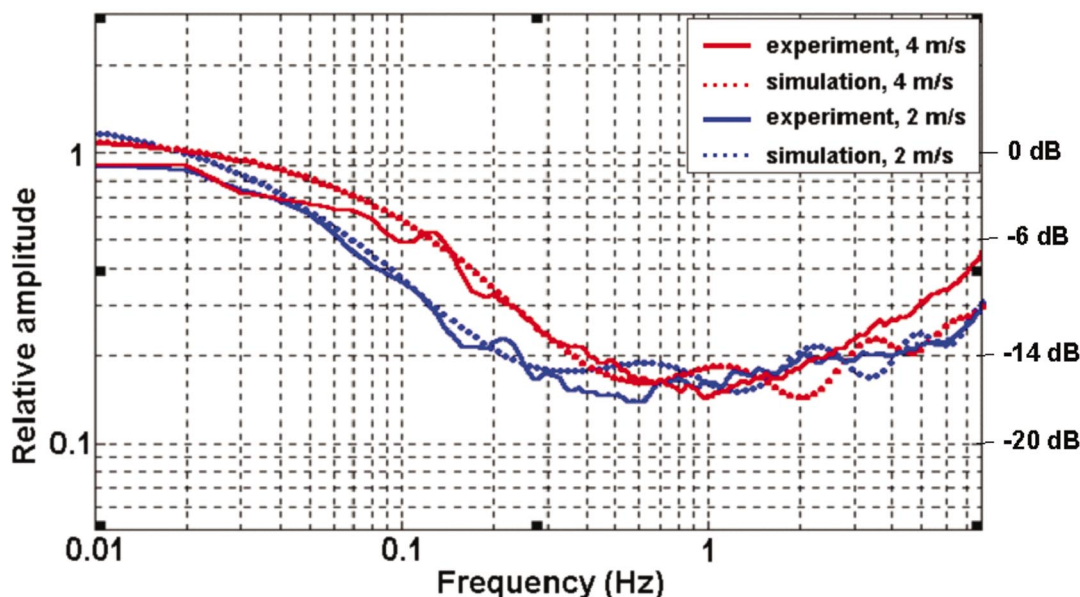


FIG. 8. Simulations and measurements of noise reduction for the 16 m system, for wind speeds of 2 and 4 m/s.

The comparison is performed between the spectra amplitudes of $y(t)$ and one $w_i(t)$.

A. Effect of the air inlets

With a diameter of 15 mm, the air inlets are essentially open. The inlets are covered by a protective 1 mm mesh wire-screen shield to prevent the entry of foreign bodies likely to block the pipes. Moreover, comparative noise measurements with and without grid screens show a significant wind-noise reduction above 5 Hz [Fig. 7(a)]. To take into account the screen effect, the noise spectrum is weighted using the following approximation function derived from measurements [Fig. 7(b)]:

$$P(f) = 0.9661 + 0.8443 \cdot f - 0.2861 \cdot f^2 + 0.0470 \cdot f^3 - 0.0036 \cdot f^4 + 0.0001 \cdot f^5. \quad (6)$$

The noise reduction simulated as a function of frequency is given by

$$E(f) = \frac{P(f) \cdot |Y(f)|}{|W(f)|}, \quad (7)$$

where $Y(f)$ and $W(f)$ are the Fourier transform of $y(t)$ and $w_i(t)$, respectively.

B. Validation experiments

The validation of the acoustic modeling of the complete system is carried out in our facilities located near Paris. The following experimental protocol is applied.

- (i) The experiment is located on a clear ground. To ensure the same noise level conditions (boundary layer effects), the air entries to the open sensor are located at the same height as the air inlets of the noise reduction system (<0.2 m). Accordingly, the sensors must be subjected to the same wind conditions (a distance of separation lower than 10 m).
- (ii) A microbarometer is connected to the noise reducer.
- (iii) An open microbarometer is used as a pressure reference.
- (iv) An anemometer is used for the measurement of wind speed (height of ~ 2 m).
- (v) All measurements channels are recorded by a 24 bits DASE digitizer at a 20 Hz sample rate and transmitted to Paris in real time via satellite link.

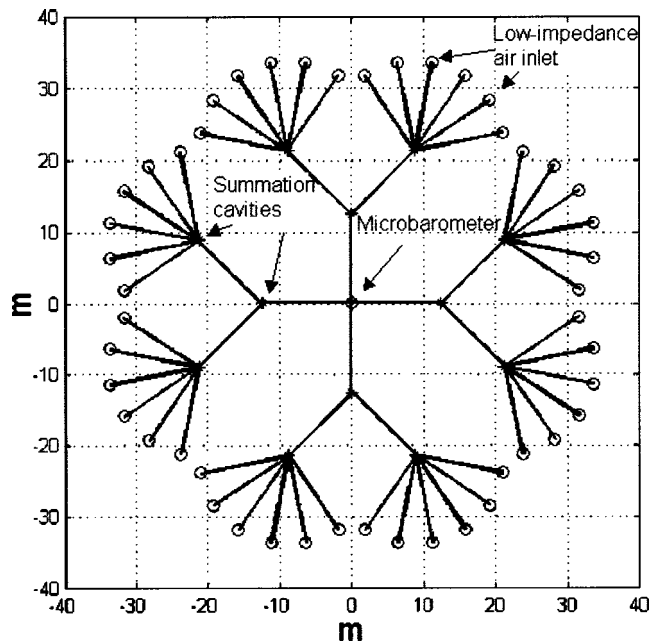


FIG. 9. Variant of the system for wind speeds higher than 5 m/s, with an average diameter of 70 m, 48 air inlets. The length of each pipe is 12.5 m.

C. Noise reduction simulated for a 16-m system

This system is optimized for wind speeds lower than 5 m/s, with: $D_m = 16$ m ($L_1 = 4$ m, $L_2 = 5$ m), 32 air inlets placed 0.2 m off the ground and 4 summation cavities (Fig. 1). Figure 8 reports the results of the noise reduction simulations compared with measurements. With a difference between simulation and measurements less than 4 dB, simulations are consistent with the measurements. The maximum noise reduction obtained is about 15 dB. The main results are the following.

- (i) For frequencies lower than 0.02 Hz, there is no noise reduction.
- (ii) A maximum noise reduction of about 15 dB is obtained at 1 Hz.
- (iii) At higher frequencies, the noise level is amplified by the first resonant frequency (Fig. 4).
- (iv) With increasing wind speed, the noise reduction pattern is shifted toward higher frequencies.

D. Noise reduction simulated for a 70 m system

When the wind speed becomes higher than 5 m/s, a 16 m system is no longer appropriate. To obtain a comparable

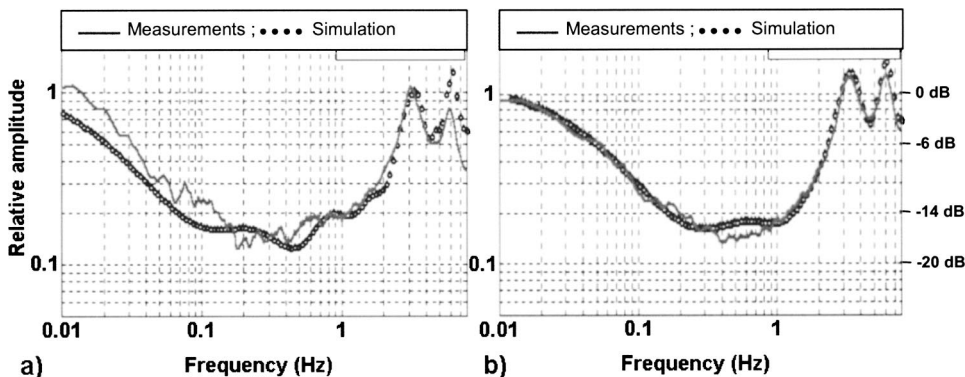


FIG. 10. Simulation and measurement of the 70 m system for wind speeds of 2 and 5 m/s. (a) Wind speed of 2 m/s. (b) Wind speed of 5 m/s.

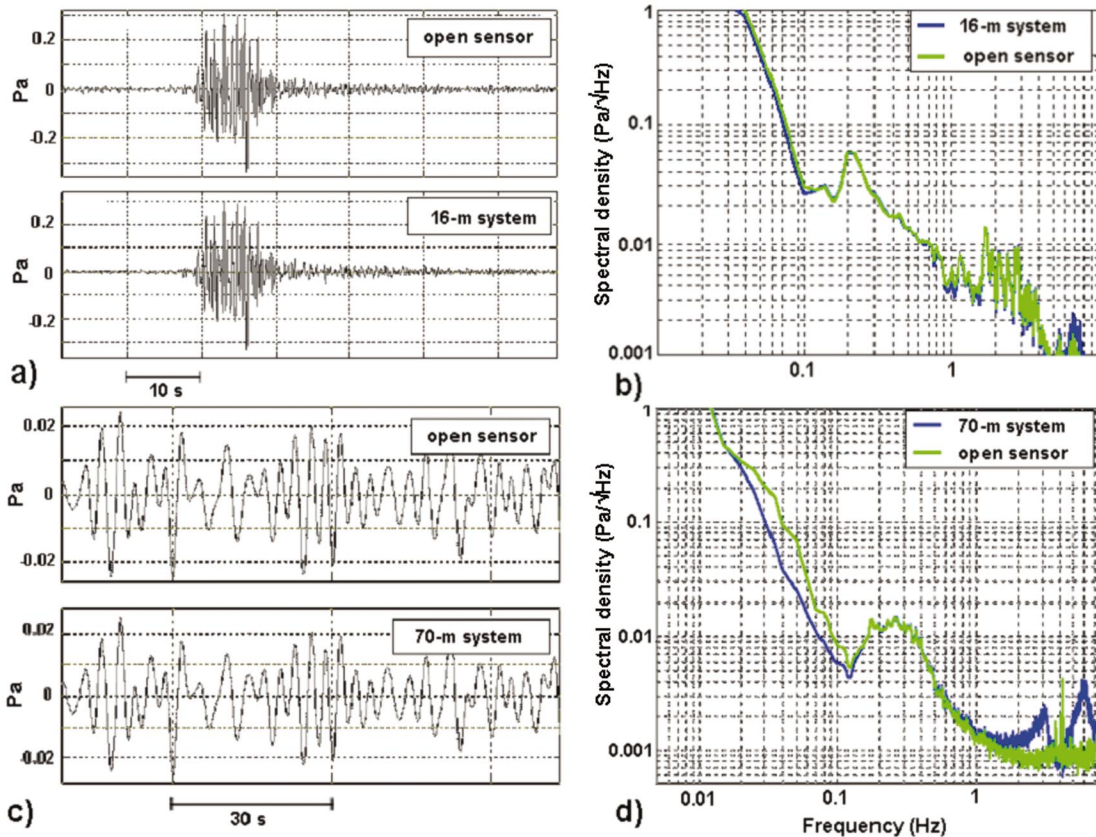


FIG. 11. Nondistortion of waveform and spectra amplitude of infrasound signals without wind. (a)–(b) Concorde signal recorded with and without the 16 m system and corresponding spectrum. (c)–(d) Microbaroms recorded with and without the 70 m system (signals filtered between 0.1 and 1 Hz) and the corresponding spectrum.

noise reduction in the frequency band [0.02–4] Hz, the diameter should be increased. Such a system is presented in Fig. 9. With an average diameter of 70 m, this arrangement comprises 48 air inlets and 12 summation cavities. Its electroacoustic model can be derived from a similar scheme presented in Fig. 2. Figure 10 shows the results of simulations of this system compared with measurements for winds of 2 and 5 m/s. The main results are the following.

- (i) For frequencies lower than 0.02 Hz, in contrast to the 16 m system (Fig. 8), the noise reduction is weak but not negligible (up to 6 dB).
- (ii) The maximum noise reduction (15 dB) is obtained between 0.1 and 1 Hz for wind speeds lower than 6 m/s. This maximal noise reduction is lower than the theoretical noise reduction provided by 48 air inlets (17 dB) (Daniels, 1959).
- (iii) The amplitude and the frequency of the resonances are well predicted. The first resonant peak is introduced by the longer elementary pipe (Hedlin *et al.*, 2003). This tends to limit the frequency band of large-sized systems to frequencies lower than 1 Hz (Grover, 1971; Noel and Whitaker, 1991).
- (iv) Compared with the 16-m system, the noise reduction pattern is shifted toward the low frequencies for the same wind speeds.

E. Effect on the signal

Figures 11(a) and 11(b) represent a signal generated by

the Concorde sonic boom recorded with and without the 16-m system without wind. These signals, consistently recorded by an array of microbarographs in France, are associated with the daily supersonic Concorde flights between North America and Europe (Le Pichon, 2002). The measured signal amplitudes with and without the noise reducer are identical (~ 0.6 Pa peak-to-peak). Figures 11(c) and 11(d) represent the signal of microbaroms from the Atlantic Ocean

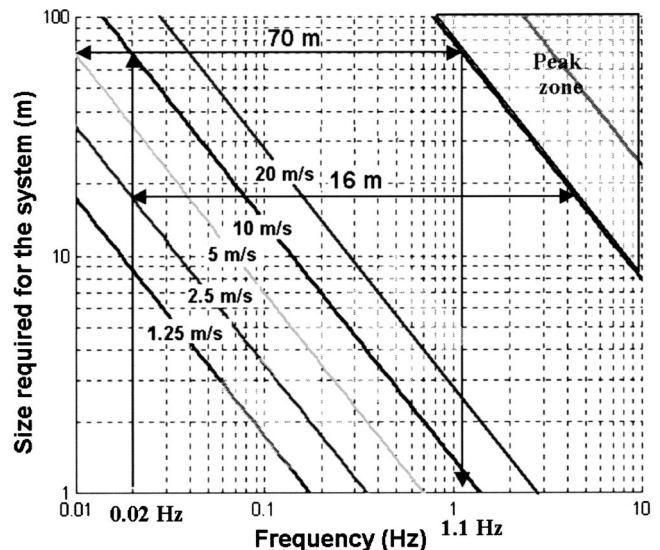


FIG. 12. Dimensioning diagram for a wind noise reducer as a function of the effective band and the average wind speed.

recorded with and without the 70-m system. These two signals have an identical time sequence and spectra amplitude. Since the level of noise is significantly reduced by these systems, an increase in the signal-to-noise ratio for infrasound measurements is expected.

F. Discussion

Following the simulations, it appears that, for a circular repartition of a finite number of inlets, we have the following.

- (i) The minimal frequency f_1 at which the noise reduction starts depends both on the wind speed V and the size of the system D_m (Figs. 8 and 10) and can be approximated by the relation $f_1 = 0.14 V/D_m$. The frequency band of noise reduction is shifted toward higher frequencies with a decreasing size of the system or increasing wind speed.
- (ii) The frequency band is limited by resonance whose frequency f_0 depends only on the size of the system [Eq. (1)].
- (iii) Because of resonance, the theoretical noise reduction provided by a large number of air inlets cannot be attained. This is particularly noticeable in the case of strong winds, where the maximal noise reduction below the resonance is lower than the theoretical noise reduction.
- (iv) The optimal number of air inlets and the size of the system can be established as a function of the average wind speed and the effective frequency band.
- (v) By removing resonances and using a large number of inlets, the noise reduction could be improved in high frequencies.

Items (i) and (ii) give the effective frequency band of noise reduction, according to the average wind speed and the outer diameter. The diagram presented by Fig. 12 dimensions the noise reduction system. Inclined lines are plotted for various wind speeds, and the figure gives the frequency of onset of noise reduction f_1 according to the array size. The appearance of acoustic resonant peaks is represented as a zone on the top right-hand corner. For a given wind speed, the minimal size of the system is given by the intersection between the minimal frequency f_1 and the lines of equal

values of wind speed. For example, for a system subjected to a wind of 10 m/s, a diameter of 70 m provides an onset of noise reduction at 0.02 Hz. The horizontal projection towards the right gives the upper frequency of the effective band (1.1 Hz). A 16-m diameter system is efficient from 0.02 to 4 Hz, for wind speeds lower than 2.5 m/s. For higher wind speeds, the noise reduction is limited to the high-frequency band (for a wind speed of 10 m/s, the effective band is [0.08–4] Hz).

V. CONCLUSION

The development of infrasound equipment in the framework of the CTBT requires the study of robust and effective systems for the reduction of acoustic noise generated by the wind. The design of a system using low-impedance acoustic elements has been investigated. The optimization of these systems has been carried out by numerical simulations based on electroacoustic modeling and the use of an appropriate wind-noise model. The noise reduction has been first simulated and then validated by experiments. The results obtained allowed us to optimize a system efficient in the [0.02–4] Hz band. The diameter of the system, as well as the maximum noise reduction attainable for a given frequency band and wind speed, have been determined. It has been shown that the low onset frequency of noise reduction depends only on the size of the system and increases with wind speed. The resonances inside the pipes limit the effective bandwidth in both frequency response and noise reduction. The noise reduction obtained with an infinite number of air inlets cannot theoretically exceed 25 dB for low wind speed and decreases with increasing wind speed. A maximum noise reduction of the order of 20 dB is obtained with about 100 air inlets. In order to enhance the performances of these systems, the resonances need to be reduced or shifted. One way would be to use purely acoustic systems based on impedance matching by means of capillaries (Christie, 2001; Hedlin, 2001).

ACKNOWLEDGMENTS

The authors gratefully acknowledge Dr. M. Hedlin for his interest in this field and for the many helpful discussions we had during the completion of this work.

APPENDIX A: ELEMENTS OF ELECTROACOUSTIC MODEL (Fig. 2)

Elements	Model	Values
Air inlet: <i>Radiation impedance</i>	From the radiation impedances (Beranek, 1986): $Z_{ar} = R_{apr} + j\omega \cdot M_{apr}$	$a < 10$ mm, $f = 10$ Hz $\rho = 1.225$ kg/m ³ $ Z_c > 1300$ k Ω A ^a $R_{apr} = 52E - 6$ Ω A ^a $M_{apr} = 0.26$ kg/m ⁴
a : radius of measurement port ρ : air density k : acoustic wave number c : sound speed in air	$R_{apr} = Z_c \frac{(ka)^2}{1 + (ka)^2}$ $M_{apr} = Z_c \frac{ka}{1 + (ka)^2}$ with: $Z_c = \frac{\rho c}{\pi a^2}$, $k = \frac{2\pi f}{c}$	$\Rightarrow R_{apr} \ll Z_c $
Pipes 1 and 2: <i>Dissipative transmission line</i>	$Z_{at11} = Z_c \cdot \tanh\left(\frac{\Gamma \cdot l}{2}\right)$	$\eta = 18.6 \times 10^{-6}$ (Pa s) at 20 °C.

APPENDIX A: (continued)

Elements	Model	Values
<i>l</i> : length of pipes η : viscosity coefficient of air <i>S</i> : cross section of the pipe	$Z_{\text{atli2}} = \frac{Z_c}{\Gamma \cdot l}$ $Z_c = R_0 \cdot [(1 + 0.369 \cdot r_v^{-1}) - j \cdot 0.369 \cdot r_v^{-1}]$ $\Gamma = \mathbf{k} \cdot [1.045 \cdot r_v^{-1} + j \cdot (1 + 1.045 \cdot r_v^{-1})]$ <p>with $R_0 = \frac{\rho \cdot c}{S}$, $r_v = \sqrt{\frac{2 \cdot \rho \cdot l \cdot S}{\eta}} \geq 1$</p>	These broad tube approximations (Keefe, 1984) are close to exact wave parameters (Benade, 1968).
Summation cavity: Acoustic compliance V_s : cavity volume	$C_{\text{avs}} = \frac{V_s}{\rho \cdot c^2}$	$V_s = 0.3 \times 10^{-3} \text{ m}^3$
Microbarometer		MB2000 type
Inlet pipes: Mass + losses l_p : length of inlets r_p : radius of inlets	$R_{\text{ap}} = \frac{8 \cdot \eta \cdot l_p}{\pi \cdot r_p^4}, \quad M_{\text{ap}} = \frac{\rho \cdot l_p}{\pi \cdot r_p^2}$	$l_p = 30 \text{ mm}, r_p = 5 \text{ mm}$
Inner volume: Acoustic compliance V_{ab} : volume of cavity and losses R_{ab}	$C_{ab} = \frac{V_{ab}}{\rho \cdot c^2}, \quad R_{ab}$	$V_{ab} = 0.6 \times 10^{-3} \text{ m}^3$ $C_{ab} = 4.33 \times 10^{-6} \text{ m}^3 \text{ Pa}$ $R_{ab} = 500 \Omega^{\text{a}}$
Sensitive bellows: Mass spring oscillator R_{ad} , M_{ad} , C_{ad} : sensor bellows	$Z_{\text{ad}} = R_{\text{ad}} + j\omega \cdot M_{\text{ad}} + \frac{1}{j\omega \cdot C_{\text{ad}}}$	$C_{\text{ad}} = 5.1 \times 10^{-11} \text{ m}^3/\text{Pa}$ $M_{\text{ad}} = 2.85 \times 10^4 \text{ kg/m}^4$ $R_{\text{ad}} = 2.62 \times 10^5 \Omega^{\text{a}}$

^a Ω A denotes an MKS acoustic ohm in Pa s/m³.

APPENDIX B: CALCULATION OF TRANSFER FUNCTIONS BY THE MATRIX METHOD

The diagram for a node *k* is shown by Fig. 13(a). The acoustic flow rate between node *k* and node *j* is written as

$$qv_{(j,k)} = \frac{N_j - N_k}{Z_{(j,k)}}, \tag{B1}$$

with $\sum_j qv_{(j,k)} = 0$ at node *k*, it follows that

$$\sum_j \frac{N_j}{Z_{(j,k)}} - N_k \cdot \sum_j \frac{1}{Z_{(j,k)}} = 0. \tag{B2}$$

If an air inlet is connected to node *k* by a pipe, the impedance is modeled by Ze_k (formed by M_{apr} , R_{apr} , and Z_{atli1} . Figure 13(b) presents the corresponding diagram. From Eq. (B2), we can write

$$\sum_j \frac{N_j}{Z_{(j,k)}} - N_k \cdot \left(\frac{1}{Ze_k} + \sum_j \frac{1}{Z_{(j,k)}} \right) = e_k \cdot - \frac{1}{Ze_k}. \tag{B3}$$

In order to simplify the establishment of transfer functions and perform the calculations, a matrix resolution method is used with the following:

- (i) the impedance matrix [*Z*] representing the impedances defined by Eqs. (B2) and (B3);
- (ii) the vector {*N*} representing the pressures at the nodes;
- (iii) the vector {1/*Ze*} representing the inlet admittances;
- (iv) the vector {*e*} representing the pressure applied to the air inlets.

We obtain

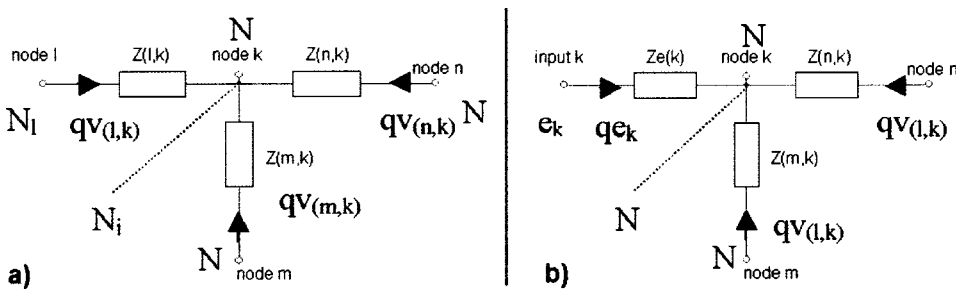


FIG. 13. Calculation of transfer functions. (a) Diagram for a node *k*. $qv_{(i,k)}$ is the volume flow rate between nodes *i* and *k*, N_j is the pressure potential at node *j* and $Z_{(i,k)}$ is the acoustic impedance between nodes *i* and *k*. (b) The calculation of impedance Ze_k .

$$\begin{bmatrix} \ddots & & \vdots & & & & \\ \cdots & & -\sum_j \frac{1}{Z_{(j,m)}} & & & & \\ & & \vdots & & \ddots & & \\ \cdots & & \frac{1}{Z_{(m,k)}} & & \cdots & -\sum_j \frac{1}{Z_{(j,k)}} & \cdots & \frac{1}{Z_{(n,k)}} & \cdots \\ & & \vdots & & \ddots & & & \vdots & \\ & & & & & & & -\sum_j \frac{1}{Z_{(j,n)}} & \cdots \\ & & & & & & & \vdots & \ddots \end{bmatrix} \cdot \begin{bmatrix} \vdots \\ N_m \\ \vdots \\ N_k \\ \vdots \\ N_m \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ -\frac{1}{Ze_m} \\ \vdots \\ -\frac{1}{Ze_k} \\ \vdots \\ -\frac{1}{Ze_n} \\ \vdots \end{bmatrix} \cdot \begin{bmatrix} \vdots \\ e_m \\ \vdots \\ e_k \\ \vdots \\ e_n \\ \vdots \end{bmatrix}$$

which has the form

$$[Z] \cdot \{N\} = - \left\{ \frac{1}{Ze} \right\} \cdot \{e\}. \quad (\text{B4})$$

The vector $\{N\}$ is then calculated at the node representing the pressure potential in the summation cavity of the microbarometer. The resolution is performed for each frequency as follows:

$$\{N(\omega)\} = [Z(\omega)]^{-1} \cdot \left\{ -\frac{e(\omega)}{Ze(\omega)} \right\}. \quad (\text{B5})$$

The transfer function linking the pressure in the sensor cavity N_p to the excitation present at air inlet e_j is obtained using a vector $\{e\}$. All the components e_k of this vector are null, except for the component corresponding to air inlet j ($e_j \neq 0$):

$$\begin{Bmatrix} \vdots \\ N_j(\omega) \\ \vdots \\ N_p(\omega) \end{Bmatrix} = [Z(\omega)]^{-1} \cdot \begin{Bmatrix} 0 \\ \vdots \\ -\frac{e_j(\omega)}{Ze_j(\omega)} \\ \vdots \\ 0 \end{Bmatrix}. \quad (\text{B6})$$

Finally, the transfer function $H_j(\omega)$ is given by

$$H_j(\omega) = \frac{N_p(\omega)}{e_j(\omega)}. \quad (\text{B7})$$

Alcoverro, B. (1998). "Acoustic filters design and experimental results," *Proceedings of the Infrasound Workshop*, Commissariat à l'Energie Atomique, Bruyères-le-Châtel, France.

Bedard, A. J. (1977). "The D-C pressure summator: Theoretical operation, experimental tests and possible practical uses," *Fluidics Quart.* **9**, 26–51.

Bedard, A. J., Whitaker, R. W., Greene, G. E., Mutschlecner, P., Nishiyama, R. T., and Davidson, M. (1992). "Measurements of pressure fluctuations near the surface of the earth," *Proceedings of the 10th Symposium on Turbulence and Diffusion*, Portland.

Benade, A. H. (1968). "On the propagation of sound waves in a cylindrical conduit," *J. Acoust. Soc. Am.* **44**, 616–623.

Beranek, L. L. (1986). *Acoustics* (American Institute of Physics, Cambridge, MA).

Burridge, R. (1971). "The acoustic of pipe arrays," *Geophys. J. R. Astron. Soc.* **26**, 53–69.

Canavero, F. G., and Einaudi, F. (1987). "Time and space variability of spectral estimates of atmospheric pressure," *J. Atmos. Sci.* **44**, 1589–1604.

Christie, D. (2001). "Wind noise reducing pipe arrays," *Proceedings of the Infrasound Workshop*, ISLA, Kailua-Kona, Hawaii.

Daniels, F. B. (1959). "Noise reduction line microphone for frequencies below 1 cps," *J. Acoust. Soc. Am.* **31**, 529–531.

Donn, W. L. (1967). "Natural infrasound of five seconds period," *Nature (London)* **215**, 1469–1470.

Einaudi, F., Bedard, A. J., and Finnigan, J. J. (1989). "A climatology of gravity waves and other coherent disturbances at the Boulder Atmospheric Observatory during March–April 1984," *J. Atmos. Sci.* **46**, 303–329.

Eliot, J. A. (1972). "Microscale pressure fluctuations measured within the lower atmospheric boundary layer," *J. Fluid Mech.* **53**, 351–383.

Gossard, E., and Munk, W. (1954). "On the gravity waves in the atmosphere," *J. Meteorol.* **11**, 259–269.

Grover, F. H. (1971). "Experimental noise reducer for an active microbarograph array," *Geophys. J. R. Astron. Soc.* **26**, 41–52.

Haak, H. W., and de Wilde, G. J. (1996). "Microbarograph systems for the infrasonic detection of nuclear explosions," *Scientific Report WR 96-06*, Royal Netherlands Meteorological Institute Seismology Division.

Hedlin, M. A. H. (2001). "Infrasonic noise reduction," *Proceedings of the Infrasound Workshop*, ISLA, Kailua-Kona, Hawaii.

Hedlin, M. A. H., Alcoverro, B., and D'Spain, G. (2003). "Evaluation of rosette infrasonic noise-reducer spatial filter," *J. Acoust. Soc. Am.* **114**, 1807–1820.

Keefe, D. H. (1984). "Acoustical wave propagation in cylindrical ducts; Transmission line parameter approximations for isothermal and nonisothermal boundary conditions," *J. Acoust. Soc. Am.* **75**, 58–62.

Kimball, B. A., and Lemon, E. R. (1970). "Spectra of air pressure fluctuations at the soil surface," *J. Geophys. Res.* **33**, 6771–6777.

Le Pichon, A., Garcés, M., Blanc, E., Barthélémy, M., and Drob, D. P. (2002). "Acoustic propagation and atmosphere characteristics derived from infrasonic waves generated by the Concorde," *J. Acoust. Soc. Am.* **111**, 629–641.

Mack, H., and Flinn, E. A. (1971). "Analysis of the spatial coherence of short-period acoustic-gravity waves in the atmosphere," *Geophys. J. R. Astron. Soc.* **26**, 255–269.

McDonald, J. A., Douze, E. J., and Herrin, E. (1971). "The structure of atmospheric turbulence and its application to the design of pipe arrays," *Geophys. J. R. Astron. Soc.* **26**, 99–109.

McDonald, J. A., and Herrin, E. (1975). "Properties of pressure fluctuations in an atmosphere boundary layer," *Boundary-Layer Meteorol.* **8**, 419–436.

Noel, S. D., and Whitaker, R. W. (1991). "Comparison of noise reduction systems," *Los Alamos National Laboratory Report LA-12008-MS UC700*.

Priestley, J. T. (1966). "Correlation studies of pressure fluctuations on the ground beneath a turbulent boundary layer," *National Bureau of Standards Report No. 8942*, U.S. Dept. of Commerce.

Wilczak, J. M., Oncley, S. P., and Bedard, A. J. (1992). "Turbulent pressure fluctuations in the atmospheric surface layers," *Proceedings of the 10th Symposium on Turbulence and Diffusion*, Portland.

Analysis of the sound field in finite length infinite baffled cylindrical ducts with vibrating walls of finite impedance

Wei Shao and Chris K. Mechefske

Department of Mechanical Engineering, Queen's University, Kingston, Ontario, Canada K7L 3N6

(Received 26 April 2004; revised 13 January 2005; accepted 14 January 2005)

This paper describes an analytical model of finite cylindrical ducts with infinite flanges. This model is used to investigate the sound radiation characteristics of the gradient coil system of a magnetic resonance imaging (MRI) scanner. The sound field in the duct satisfies both the boundary conditions at the wall and at the open ends. The vibrating cylindrical wall of the duct is assumed to be the only sound source. Different acoustic conditions for the wall (rigid and absorptive) are used in the simulations. The wave reflection phenomenon at the open ends of the finite duct is described by general radiation impedance. The analytical model is validated by the comparison with its counterpart in a commercial code based on the boundary element method (BEM). The analytical model shows significant advantages over the BEM model with better numerical efficiency and a direct relation between the design parameters and the sound field inside the duct. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1867832]

PACS numbers: 43.20.Bi, 43.20.Mv, 43.20.Rz [TDM]

Pages: 1728–1736

I. INTRODUCTION

Magnetic resonance imaging (MRI) scanners are popular medical diagnostic devices providing useful information about a patient's internal organs in a noninvasive manner. However, a disadvantage of MRI is the high level of acoustic noise generated during scanning. This noise is caused by the Lorentz forces acting on the gradient coils bound within an epoxy resin cylinder. As the gradient current switches direction the gradient windings vibrate, leading to the generation of sound waves.¹

Recently measures have been taken to control the acoustic noise generated by MRI scanners. The technique of active noise control (ANC) for the reduction of MRI noise has been studied.^{2–4} ANC was found to be effective only at relatively low frequencies. Mansfield *et al.*⁵ proposed a technique called active acoustic screening for quiet gradient coil design. While reducing the acoustic noise inside MRIs, this type of screening inevitably reduces the gradient strength as well. Yoshida *et al.*⁶ used “independent suspension” of the coil to dampen vibrations, while a “vacuum vessel enclosure” of the coil shields transmission of sound waves through the air. These methods are not suitable for all kinds of new MRI scanners and are impossible to retrofit to existing machines.

Optimizing the design of the gradient coil system could attenuate the root cause of acoustic noise. A clear understanding on the characteristics of structural vibration and acoustic radiation of the gradient coil system is necessary for the design of quiet MRI scanners. Kessels⁷ has developed a mathematical model of gradient coil structural vibrations due to the Lorentz forces acting on the windings. An acoustical model has been studied by Kuijpers.⁸ He used a baffled finite cylindrical duct with vibrating walls to describe the sound radiation of a gradient coil cylinder by assuming that only the cylinder could vibrate and thus radiate acoustic energy. This model can only be applied to the duct walls with zero acoustic admittance (or infinite acoustic impedance). It may

not always be reasonable to regard the walls as acoustically rigid. This limitation also restricts the application of this method to a gradient coil covered by sound absorption materials.

Studies on the acoustic radiation in cylindrical ducts have been reported,^{9–15} but they mainly discuss situations with simple sources (point source) or sound propagations inside ducts without vibrating walls. Models dealing with more complicated vibration conditions have also been reported.^{16,17} These models only investigate the sound field in closed cylindrical ducts (chambers).

A schematic cross section of a gradient coil system of a typical MRI scanner is shown in Fig. 1 and can be described as an axisymmetric cylinder, that is, a finite cylindrical duct with infinite baffles at the two open ends. The isocenter is the coordinate origin for the model.

Due to wave reflection at the openings of a duct, the acoustic radiation impedances at the open ends should be considered as a boundary condition for the analytical model. The radiation impedances at the open end of a finite cylindrical duct with an infinite flange were studied by Zorumski.¹¹ His results showed that the acoustical conditions of the duct wall have a significant influence on the radiation impedances.

In this paper, the radiation impedances are calculated for a finite cylindrical duct with rigid and absorptive walls. Based on these results, the inside sound field generated by the vibrating wall is simulated. The variations of the pressure levels caused by applying different acoustic admittances to the wall are also shown and discussed. The results obtained from the analytical models are compared with the data calculated using the commercial software SYSNOISE, which is based on the boundary element method (BEM). Compared with the BEM, the most important feature of analytical methods is that they can generally show the dominant parameters for the modeled problems more directly and, therefore, yield

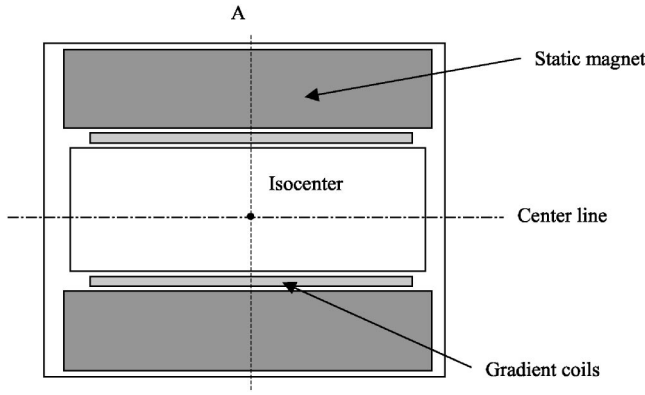


FIG. 1. Schematic cross-sectional view of an MRI scanner.

more physical insight into the problem. In addition, they are much more computationally efficient.

II. THEORY

An analytical model for the sound field generated by the structural vibration of the wall of a finite cylindrical duct with finite acoustic impedance will be derived in this section. First, the fundamental theory of duct acoustics will be reviewed briefly. Then, the boundary problem will be solved by making the acoustical model satisfy both boundary conditions at the duct wall (vibrating and with a finite acoustic impedance) and the two open ends (radiation impedances).

A. Sound field of a infinite cylindrical duct

The sound field inside an infinite cylindrical duct can be expressed as a sum of modal solutions (a time factor $e^{i\omega t}$ is understood throughout this paper),

$$p(r, \theta, x) = \sum_{m=-\infty}^{\infty} \sum_{n=1}^{\infty} J_m(\alpha_r^{mn} r) e^{-im\theta} [A_{mn} e^{-i\alpha_x^{mn} x} + B_{mn} e^{i\alpha_x^{mn} x}], \quad (1)$$

where x is the coordinate in the axial direction of the duct, r is in the radial direction, and θ is in the circumferential direction (see Fig. 2). $J_m(\alpha_r^{mn} r)$ are the Bessel functions of the first kind with circumferential order m . Here n is the radial mode number. The parameters α_x and α_r are the wave num-

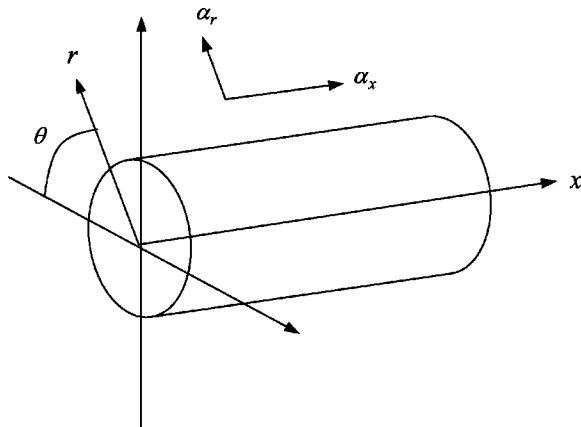


FIG. 2. Cylindrical polar coordinates.

bers in the axial and the radial direction, respectively. A and B are the modal coefficients of the forward- and backward-propagating acoustic wave modes, respectively. For the rigid wall case, the boundary condition can be expressed as

$$J'_m(\alpha_r^{mn} a) = 0. \quad (2)$$

For the wall with a finite acoustic impedance, the boundary condition should satisfy the following equation:¹⁸

$$\alpha_r^{mn} J'_m(\alpha_r^{mn} a) = i\beta k J_m(\alpha_r^{mn} a), \quad (3)$$

where J'_m is the derivative of J_m with respect to its argument; a is the radius of the duct and β is a specific acoustic admittance defined as $\beta = \xi - i\sigma = \rho c/Z$. It should be noted that $Z = p/u_n$, where Z is the acoustic impedance, p is the pressure on the surface, and u_n is the normal surface velocity. The parameters ρ and c are the density and speed of sound in the media, respectively, and k is the free field wave number.

The velocity distribution of the vibrating wall can be written in the form of a Fourier series,

$$\begin{aligned} u_r(a, \theta, x) &= U(\theta, x) \\ &= \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} e^{im\theta} U_m(x) \\ &= \frac{1}{4\pi^2} \sum_{m=-\infty}^{\infty} e^{-im\theta} \int_{-\infty}^{\infty} \tilde{U}_m e^{-i\alpha_x x} d\alpha_x. \end{aligned} \quad (4)$$

Assuming the length of the vibrating wall is $2L$, the Fourier coefficients, \tilde{U}_m , are defined by

$$\tilde{U}_m = \int_0^{2\pi} e^{im\theta} d\theta \int_{-L}^L U(\theta, x) e^{i\alpha_x x} dx. \quad (5)$$

By considering the acoustic impedance of the wall, the sound pressure should satisfy the combined boundary condition at the wall,

$$\begin{aligned} \left. \frac{\partial p}{\partial r} \right|_{r=a} &= i\rho\omega u_r(a, \theta, x) + ik\beta p \\ &= \frac{i\rho ck}{4\pi^2} \sum_{m=-\infty}^{\infty} e^{-im\theta} \int_{-\infty}^{\infty} \tilde{U}_m e^{-i\alpha_x x} d\alpha_x + ik\beta p. \end{aligned} \quad (6)$$

In Eq. (6) the first term indicates the effect of the surface velocity distribution; the second term is the influence of the finite impedance.

The sound pressure can also be expressed by the spatial Fourier transformation in the axial direction and the Fourier series in the circumferential azimuth, θ , of the pressure mode. Therefore, a solution for the sound field is

$$p(r, \theta, x) = \sum_{m=-\infty}^{\infty} e^{-im\theta} \int_{-\infty}^{\infty} A_m J_m(\alpha_r r) e^{-i\alpha_x x} d\alpha_x, \quad (7)$$

where $\alpha_r^2 = k^2 - \alpha_x^2$, and A_m are the coefficients. At a particular frequency and for a certain acoustic mode, the axial wave number $\alpha_x^{mn} = \sqrt{k^2 - \alpha_r^{mn}} = 0$. This is called the cutoff frequency below which the particular modes cannot propagate

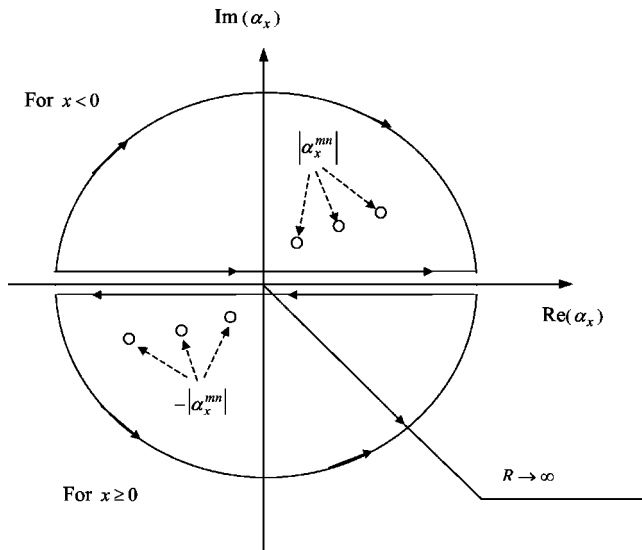


FIG. 3. The residue integration contour.

freely in the duct. Those axial modes above the cutoff frequency are called cut-on modes.

Bringing the above equation into Eq. (6) to solve for the boundary condition, the following expression can be obtained:

$$\alpha_r A_m J'_m(\alpha_r r) = \frac{i\rho ck}{4\pi^2} \tilde{U}_m + ik\beta A_m J_m(\alpha_r r). \quad (8)$$

The sound pressure can now be represented by replacing A_m in Eq. (7),

$$p(r, \theta, x) = \frac{i\rho ck}{4\pi^2} \sum_{m=-\infty}^{\infty} e^{-im\theta} \int_{-\infty}^{\infty} \frac{\tilde{U}_m}{\alpha_r J'_m(\alpha_r a) - ik\beta J_m(\alpha_r a)} \times J_m(\alpha_r r) e^{-i\alpha_x x} d\alpha_x. \quad (9)$$

This equation can be solved by the residue integration method. An integration contour is chosen to enclose the lower half-plane when $x \geq 0$ and enclose the upper half-plane, where $x < 0$, as shown in Fig. 3. The poles, α_x^{mn} , inside the contour can be calculated from

$$\alpha_x^{mn} = \pm \sqrt{k - \alpha_r^{mn}}, \quad (10)$$

where α_r^{mn} are solved by rewriting Eq. (3) as $\alpha_r J'_m(\alpha_r r) - ik\beta J_m(\alpha_r r) = 0$. When $x \geq 0$, only the poles α_x^{mn}

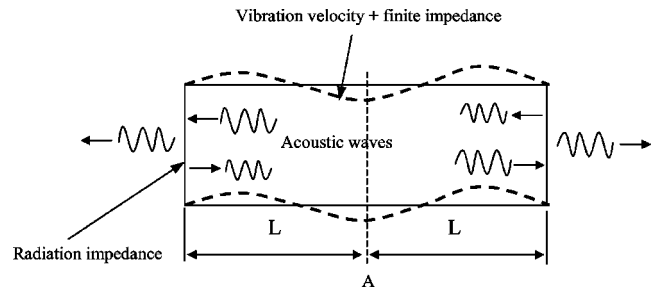


FIG. 4. Boundary conditions for the finite cylinder. (The sound field should satisfy both the impedance and velocity boundary conditions at the duct wall and the radiation impedance boundary condition at the open ends.)

$= -\sqrt{k - \alpha_r^{mn}}$ are inside the contour. Conversely, the poles $\alpha_x^{mn} = \sqrt{k - \alpha_r^{mn}}$ are inside the integration contour for $x < 0$. Thus the integral in Eq. (9) can be calculated by the sum of the residues at these poles (see Appendix A) and it can be rewritten as

$$p(r, \theta, x) = \frac{\rho ck}{2\pi} \sum_{m=-\infty}^{\infty} e^{-im\theta} \sum_{n=1}^{\infty} \frac{(\alpha_r^{mn})^2}{\alpha_x^{mn}} \times \frac{J_m(\alpha_r^{mn} r)}{(\alpha_r^{mn})^2 a J'_m(\alpha_r^{mn} a) + ik\beta J_m(\alpha_r^{mn} a)(1 - ika\beta)} \times \int_{-L}^L U_m(x') e^{-i\alpha_x^{mn}|x-x'|} dx' \dots, \quad (11)$$

where $\text{Re}(\alpha_x^{mn})$ and $\text{Im}(\alpha_x^{mn}) \geq 0$ for $x < 0$; $\text{Re}(\alpha_x^{mn})$ and $\text{Im}(\alpha_x^{mn}) < 0$ for $x \geq 0$.

B. Sound field of a finite cylindrical duct

The expression of the sound pressure level, $P(r, \theta, x)$, for a finite duct must satisfy the boundary at the wall caused by the structural vibration and finite acoustic impedance. The expression also needs to agree with radiation impedances at the duct open ends (see Fig. 4). Therefore, the reflection at the open ends should be taken into consideration and the total pressure needs to be expressed as a sum of the above solution satisfying the boundary condition at the wall and a general solution. The total solution that satisfies the boundary conditions at both the wall and the open ends can be represented as

$$p(r, \theta, x) = p'(r, \theta, x) + p''(r, \theta, x), \quad (12)$$

where

$$p'(r, \theta, x) = \frac{\rho ck}{2\pi} \sum_{m=-\infty}^{\infty} e^{-im\theta} \sum_{n=1}^{\infty} \frac{(\alpha_r^{mn})^2}{\alpha_x^{mn}} \frac{J_m(\alpha_r^{mn} r)}{(\alpha_r^{mn})^2 a J'_m(\alpha_r^{mn} a) + ik\beta J_m(\alpha_r^{mn} a)(1 - ika\beta)} \int_{-L}^L U_m(x') e^{-i\alpha_x^{mn}|x-x'|} dx', \quad (13)$$

$$p''(r, \theta, x) = \sum_{m=-\infty}^{\infty} e^{-im\theta} \sum_{n=1}^{\infty} B_{mn} J_m(\alpha_r^{mn} r) (e^{-i\alpha_x^{mn} x} + e^{i\alpha_x^{mn} x}). \quad (14)$$

In these expressions, B_{mn} are the coefficients for the general solution and it is assumed that they are the same for the forward-propagating ($e^{-i\alpha_x^{mn}x}$) and backward-propagating ($e^{i\alpha_x^{mn}x}$) acoustic wave modes because the duct is symmetric about the center plane A (see Fig. 4). In fact, the velocity distribution of the duct wall is also symmetric about the center plane due to the symmetry of the gradient coil windings. Therefore, the total solution for the sound field satisfies the boundary condition:

$$\frac{\partial p}{\partial r} = \frac{\partial p'}{\partial r} + \frac{\partial p''}{\partial r} = i\rho\omega u_r(a, \theta, x) + ik\beta p. \quad (15)$$

The expression for the total pressure can be rearranged as

$$p(r, \theta, x) = \sum_{m=-\infty}^{\infty} e^{-im\theta} \sum_{n=1}^{\infty} J_m(\alpha_r^{mn}r) [(G'_{mn}(x) + B_{mn}) \times e^{-i\alpha_x^{mn}x} + (G''_{mn}(x) + B_{mn})e^{i\alpha_x^{mn}x}] \quad (16)$$

with

$$G'_{mn}(x) = \frac{\rho ck}{2\pi} \frac{(\alpha_r^{mn})^2}{\alpha_x^{mn}[(\alpha_r^{mn})^2 a J_m''(\alpha_r^{mn}a) + ik\beta J_m(\alpha_r^{mn}a)(1 - ik\alpha\beta)]} \int_{-L}^x U_m(x') e^{i\alpha_x^{mn}x'} dx', \quad (17)$$

$$G''_{mn}(x) = \frac{\rho ck}{2\pi} \frac{(\alpha_r^{mn})^2}{\alpha_x^{mn}[(\alpha_r^{mn})^2 a J_m''(\alpha_r^{mn}a) + ik\beta J_m(\alpha_r^{mn}a)(1 - ik\alpha\beta)]} \int_{-x}^L U_m(x') e^{i\alpha_x^{mn}x'} dx', \quad (18)$$

The pressure at an open end (suppose the right side, $x=L$) is

$$p(r, \theta, L) = \sum_{m=-\infty}^{\infty} e^{-im\theta} \sum_{n=1}^{\infty} [(G_{mn} + B_{mn})e^{-i\alpha_x^{mn}L} + B_{mn}e^{i\alpha_x^{mn}L}] J_m(\alpha_r^{mn}r). \quad (19)$$

with

$$G_{mn}(x) = \frac{\rho ck}{2\pi} \frac{(\alpha_r^{mn})^2}{\alpha_x^{mn}[(\alpha_r^{mn})^2 a J_m''(\alpha_r^{mn}a) + ik\beta J_m(\alpha_r^{mn}a)(1 - ik\alpha\beta)]} \int_{-L}^L U_m(x') e^{i\alpha_x^{mn}x'} dx'. \quad (20)$$

The corresponding axial velocity is

$$u_x(r, \theta, L) = \frac{1}{\rho c} \sum_{m=-\infty}^{\infty} e^{-im\theta} \sum_{n=1}^{\infty} \frac{\alpha_x^{mn}}{k} [B_{mn}e^{i\alpha_x^{mn}L} - (G_{mn} + B_{mn})e^{-i\alpha_x^{mn}L}] J_m(\alpha_r^{mn}r). \quad (21)$$

The acoustic pressure and velocity amplitudes at the open ends of the duct can be expressed in terms of the acoustic modes in radial r and circumferential θ directions as¹¹

$$p(r, \theta, x) = \sum_{m=-\infty}^{\infty} e^{-im\theta} \sum_{n=1}^{\infty} P_{mn} J_m(\alpha_r^{mn}r), \quad (22)$$

$$u_x(r, \theta, x) = \frac{1}{\rho c} \sum_{m=-\infty}^{\infty} e^{-im\theta} \sum_{n=1}^{\infty} V_{mn} J_m(\alpha_r^{mn}r), \quad (23)$$

where P_{mn} and V_{mn} are the modal coefficients for the pressure and velocity, respectively.

Comparing Eqs. (22) and (23) to Eqs. (19) and (21), respectively, the amplitudes of modal pressure P_{mn} and velocity V_{mv} are

$$P_{mn} = (G_{mn} + B_{mn})e^{-i\alpha_x^{mn}L} + B_{mn}e^{i\alpha_x^{mn}L} = B_{mn}(e^{i\alpha_x^{mn}L} + e^{-i\alpha_x^{mn}L}) + G_{mn}e^{-i\alpha_x^{mn}L}, \quad (24)$$

$$V_{mv} = \frac{\alpha_x^{mn}}{k} [B_{mn}e^{i\alpha_x^{mn}L} - (G_{mn} + B_{mn})e^{-i\alpha_x^{mn}L}] = \frac{\alpha_x^{mn}}{k} [B_{mn}(e^{i\alpha_x^{mn}L} - e^{-i\alpha_x^{mn}L}) - G_{mn}e^{-i\alpha_x^{mn}L}]. \quad (25)$$

The relationship between the modal pressure and velocity amplitudes can be expressed by the radiation impedance, Z , at the open end with an infinite flange,

$$P_{mn} = \sum_{l=1}^{\infty} Z_{mnl} V_{ml}, \quad (26)$$

where l and n are the orders of radial incident and reflected modes, respectively. Referring to Appendix B, the radiation impedance, Z , can be calculated using

$$Z_{mnl} = \frac{1}{W} \int_0^{\pi/2} \frac{\sin^3 \phi J_m'^2(ka \sin \phi) - 2i\beta \sin^2 \phi J_m(ka \sin \phi) J_m'(ka \sin \phi) - \beta^2 \sin \phi J_m^2(ka \sin \phi)}{(\sin^2 \phi - (\alpha_r^{mn})^2/k^2)(\sin^2 \phi - (\alpha_r^{ml})^2/k^2)} d\phi$$

$$+ \frac{i}{W} \int_0^\infty \frac{\cosh^3 \xi J_m'^2(ka \cosh \xi) - 2i\beta \cosh^2 \xi J_m(ka \cosh \xi) J_m'(ka \cosh \xi) - \beta^2 \cosh \xi J_m^2(ka \cosh \xi)}{(\cosh^2 \phi - (\alpha_r^{mn})/k^2)(\cosh^2 \phi - (\alpha_r^{ml})^2/k^2)} d\xi. \quad (27)$$

Bringing the above two equations for P_{mn} and V_{mv} into Eq. (26) gives

$$B_{mn}(e^{i\alpha_x^{mn}L} + e^{-i\alpha_x^{mn}L}) + G_{mn}e^{-i\alpha_x^{mn}L}$$

$$= \sum_{v=1}^{\infty} Z_{mnv} \frac{\alpha_x^{mn}}{k} [B_{mn}(e^{i\alpha_{xm}^n L} - e^{-i\alpha_x^{mn}L}) - G_{mn}e^{-i\alpha_x^{mn}L}]. \quad (28)$$

For a fixed order, m , the above equation can be expressed in matrix form as

$$\mathbf{g}_m \mathbf{E}_m + \mathbf{E}_m' \mathbf{b}_m = \mathbf{Z}_m (\mathbf{K}_m \mathbf{E}_m'' \mathbf{b}_m - \mathbf{K}_m \mathbf{g}_m)$$

for $m=0, \pm 1, \pm 2, \dots$. (29)

The detailed expressions of the matrices in Eq. (29) are presented in Appendix C. The coefficient matrix \mathbf{b}_m can be solved using the following equation,

$$[\mathbf{E}_m + \mathbf{Z}_m \mathbf{K}_m] \mathbf{g}_m = [\mathbf{Z}_m \mathbf{K}_m \mathbf{E}_m'' - \mathbf{E}_m'] \mathbf{b}_m. \quad (30)$$

Once \mathbf{b}_m is known, these values can be brought into Eq. (16) and the sound field pressure, p , can be calculated.

III. NUMERICAL RESULTS AND DISCUSSION

To verify the validity of the above mathematical model, the sound pressures at various points within the sound field inside a finite cylindrical duct were calculated. The radii of the gradient coils of MRI scanners are normally about 0.3 m. The radius of the gradient coils for the 4 Tesla, Varian/Siemens, Unity INOVA whole-body MRI system, used in previously reported experimental studies,⁴ is 0.34 m (1.2 m in length). Therefore, a cylindrical duct with 0.3-m radius and 1.2-m length was used for the simulations and different wall acoustic impedances (rigid, $\beta=0.1$ and $\beta=0.3$) were applied. The analytical results are compared with the data calculated using the commercial code LMS SYSNOISE, which is based on the boundary element method (BEM).

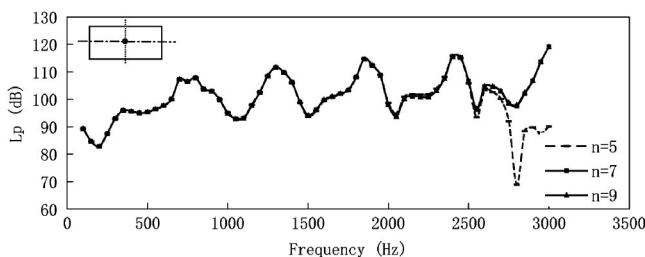


FIG. 5. The sound pressure level, L_p , at the isocenter calculated with different truncated orders, $n=5, 7$ and 9 (the wall admittance $\beta=0.1$).

The velocity distribution on the duct wall can normally be expressed by a sum of trigonometric functions in the circumferential and axial directions, such as

$$U = \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} U_A \cos(m\theta) \cos(n\pi x/2L). \quad (31)$$

The amplitude, $U_A=0.001$ m/s, is arbitrarily chosen. The parameter L is the length of the duct; $m=0$ and $n=2$, and $m=1$ and $n=3$, are used in the calculations. Thus the velocity distribution can be written as

$$U = 0.001 [\cos(2\pi x/1.2) + \cos(\theta) \cos(3\pi x/1.2)]. \quad (32)$$

Equation (13) is used for calculating the pressure. It is apparent, because of the Bessel functions [see Eq. (33)], that just the first term, $0.001 \cos(2\pi x/1.2)$ ($m=0$), of the duct vibration velocity contributes to the sound pressure at the center line ($r=0$):

$$J_m(\alpha_r^{mn} r)|_{r=0} = 1 \quad \text{when } m=0,$$

$$J_m(\alpha_r^{mn} r)|_{r=0} = 0 \quad \text{when } m \neq 0. \quad (33)$$

The infinite matrices for the analytical model of finite ducts are truncated. The sound pressure level, L_p (dB) (with a reference of $20 \mu\text{Pa}$), from 100 to 3000 Hz at the isocenter and the point $r=0, x=0.3$ m for the duct wall with the admittance $\beta=0.1$ calculated with different truncated orders, $n=5, 7$ and 9 , are shown in Figs. 5 and 6, respectively. It is obvious that there is a significant difference between the results calculated by order $n=5$ and the other two situations (orders $n=7$ and $n=9$) at the higher frequencies (>2400 Hz). The reason is that higher duct modes contribute little to the sound pressure below their cutoff frequency. For the analysis frequency up to 3000 Hz, there are harmonic orders from $n=1$ to $n=6$ ($m=0$) with their corresponding cutoff frequencies below 3000 Hz. (see Table I). This also holds for the order $m=1$. This suggests that infinite matrices truncated at the order $n=5$ are not accurate enough at high frequencies (>2400 Hz). However, using truncated orders $n=7$ or n

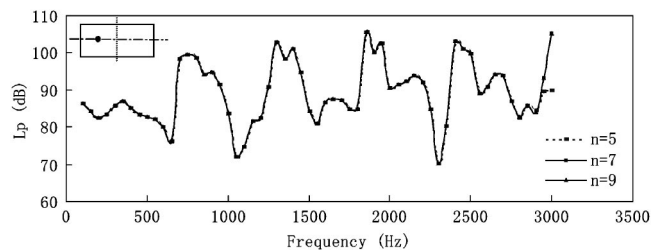


FIG. 6. The sound pressure level, L_p , at the point $r=0, x=0.3$ m calculated with different truncated orders $n=5, 7$ and 9 (the wall admittance $\beta=0.1$).

TABLE I. Cutoff frequencies, α_{cutoff} , in Hz corresponding to radial acoustic modes, α_r^{mn} . (For the rigid wall, $m=0$ and $m=1$.)

n	1	2	3	4	5	6
$m=0$	0	691	1266	1834	2402	2970
$m=1$	332	961	1540	2113	2681	3259

=9 will be acceptable for analysis of frequencies up to 3000 Hz.

It is impossible to build a duct with infinite baffles with the SYSNOISE software. However, the size of the baffles can be defined as relatively large in the BEM model for comparison to the analytical model for the finite duct with infinite flanges. For most MRI scanners, the gradient coils are actually baffled by the casing (1–1.5 m in radius) of the static magnet. Therefore, for assessing the influence of the baffle radius to acoustic response of the BEM model, the L_p at the isocenter and a particular point on the center line ($r=0, x=0.3$ m) for different baffles with radius R of 1.5, 3, and 5 m (rigid wall) are calculated and shown in Figs. 7 and 8, respectively. These results are quite close (except for a little difference at the lower frequencies) especially for the baffles of 3 and 5 m in radius. These results indicate that the BEM model with baffle 3 or 5 m validates the analytical model. The results are applicable to simulate the acoustic response of a real MRI scanner. For an analysis frequency up to 3000 Hz, at least 7400 elements are used for the BEM model to satisfy the minimum accuracy requirement.

The sound pressure level, L_p , at the isocenter for different duct wall conditions (rigid, $\beta=0.1$ and $\beta=0.3$) calculated by the BEM model (SYSNOISE) and the analytical model are shown in Fig. 9. This figure shows that the sound pressure level at the isocenter calculated by both of the methods reaches a good agreement. The pressure peaks appear around 700, 1300, 1850, 2400, and 3000 Hz (around cutoff frequencies for $m=0$, referring to Table I). This indicates that the sound energy at the isocenter of the finite cylindrical duct is concentrated around the duct cutoff frequencies. This is due to the fact that the wave reflections at the open ends reach their maximum value at the cutoff frequencies, causing acoustic resonance inside the duct. The cutoff frequencies are dependent on the geometrical dimension (radius) of the duct.

Compared with the results calculated by the analytical model for the duct with acoustic rigid wall, reductions of the sound pressure level L_p at the isocenter around cutoff frequencies are shown in Table II when absorptive boundary conditions for the duct wall are applied. The noise reduction is obvious especially at the first peak at 700 Hz and the fifth

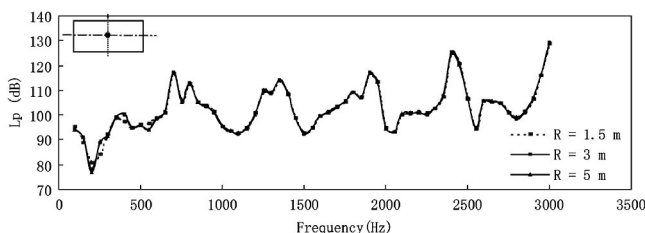


FIG. 7. The sound pressure level, L_p , at the isocenter with different baffle radii: 1.5, 3, and 5 m (rigid wall).

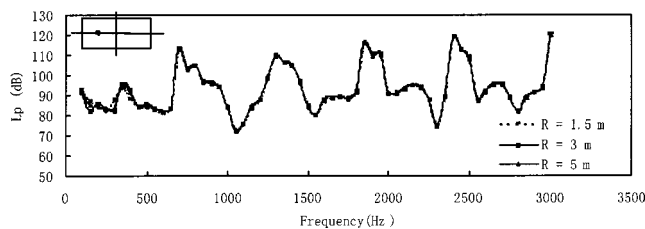


FIG. 8. The sound pressure level, L_p , at the point $r=0, x=0.3$ m with different baffle radii: 1.5, 3, and 5 m (rigid wall).

peak at 3000 Hz. These results also reveal that more noise attenuation takes place as the value of the admittance, β , increases.

Since the sound field inside the duct is symmetrical to the isocenter, the acoustical pressure distributions for half the duct ($r=0-0.2$ m, $x=0-0.6$ m) with the wall admittance of $\beta=0.1$ were calculated by the analytical model. The results from the analytical method and the differences between the analytical model and the BEM model are shown in Figs. 10–13 for the frequencies 700, 1300, 1850, and 2400 Hz, respectively. The results show that the sound fields obtained by both methods are similar except for some differences close to the open end.

IV. CONCLUSIONS

An analytical model for the acoustic radiation inside finite cylindrical ducts with infinite flanges has been presented. This model has been validated for a duct wall with acoustically rigid boundary conditions and absorptive boundary conditions. The acoustic response of the analytical model

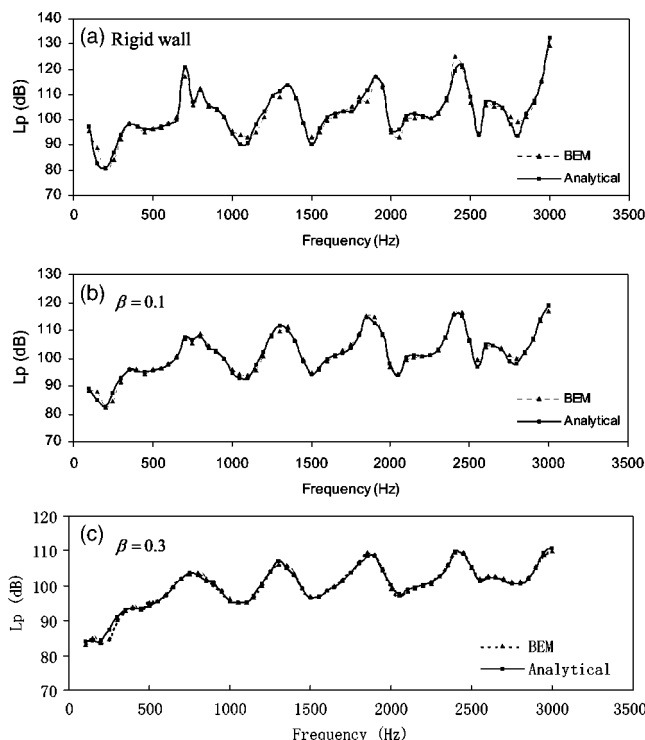


FIG. 9. The sound pressure level, L_p , at the isocenter calculated by the BEM and the analytical method for different wall conditions: (a) rigid, (b) $\beta=0.1$, and (c) $\beta=0.3$.

TABLE II. The sound pressure level, L_p (dB), reduction around the cutoff frequencies when absorptive boundary conditions are applied.

Frequency (Hz)	700	1300	1850	2400	3000
$\beta=0.1$	13	1.9	2.3	4	13.1
$\beta=0.3$	16	6.5	8	9.5	21.3

was found to be in a good agreement with the results obtained by the BEM model with LMS SYSNOISE.

When compared with the BEM model, the analytical model more directly shows the relationship of the sound field to the geometrical parameters and boundary conditions of the cylindrical duct. The observation of acoustic energy concentration around the cutoff frequencies can make it possible for MRI scanner designers to avoid designing inherently noisy gradient coils. The simulation results also show that using absorptive materials to cover the duct wall may be an effective way to reduce the noise without changing the design of the scanner. The other important benefit of using the analytical model is that reduced computer resources are required and it runs much faster than the BEM model, especially for the ducts with absorptive walls and a wide range of analysis

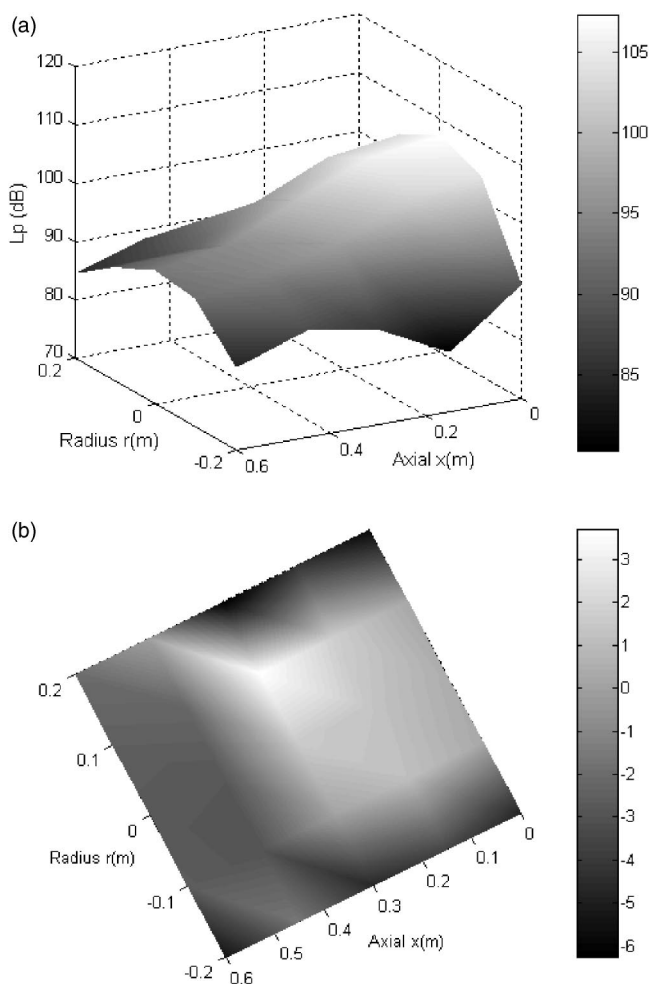


FIG. 10. The sound field (L_p) at 700 Hz ($\beta=0.1$). Axial direction from 0 to 0.6 m (the isocenter to the open end, radius from -0.2 to $+0.2$, “+” means the point in the left side to the isocenter, “-” means the right side). (a) Calculated by the analytical model. (b) The difference of results between the analytical model and the BEM model.

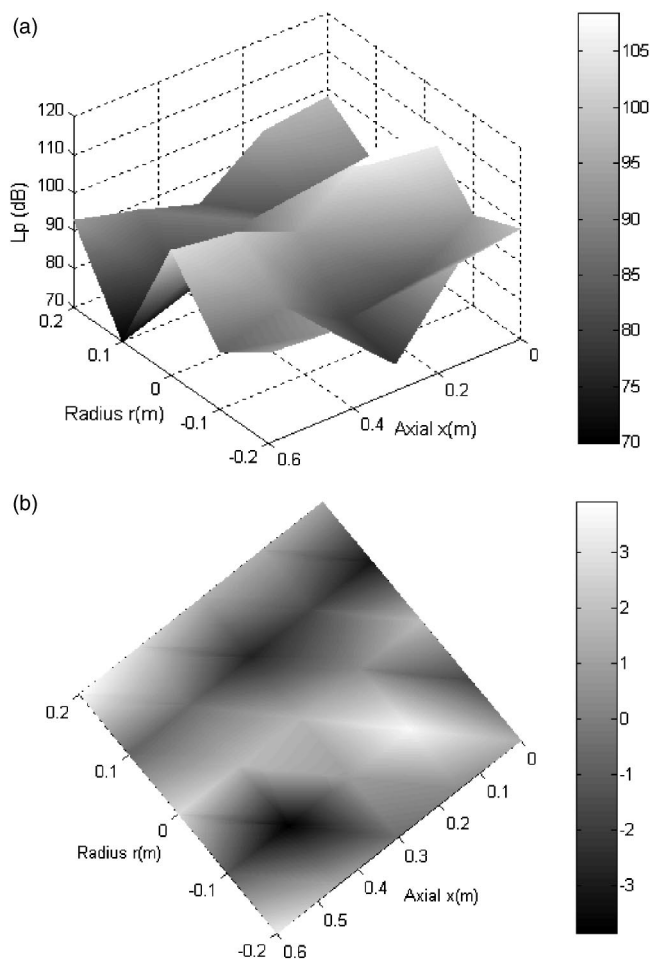


FIG. 11. The sound field (L_p) at 1300 Hz ($\beta=0.1$). (a) Calculated by the analytical model. (b) The difference of results between the analytical model and the BEM model.

frequencies. As an example, it normally took more than 2 weeks for calculating a model with absorptive walls from 100 to 3000 Hz (with a frequency resolution of 50 Hz) using SYSNOISE on a computer with an Intel P4 1.7G Hz CPU and 1.5 GB memory. The same calculation required only about 2 h for the analytical model on the same computer.

APPENDIX A: RESIDUE INTEGRATION

From the Eq. (9), the integration

$$\int_{-\infty}^{\infty} \frac{J_m(a_r r)}{\alpha_r J'_m(a_r a) - ik\beta J_m(a_r a)} e^{-i\alpha_x x} d\alpha_x$$

could be solved by the residue integration method,¹⁸

$$\int_{-\infty}^{\infty} F(\alpha_x) e^{-i\alpha_x x} d\alpha_x = \begin{cases} -2\pi i \sum \text{Re}^-(\alpha_x^{mn}), & x \geq 0, \\ 2\pi i \sum \text{Re}^+(\alpha_x^{mn}), & x < 0, \end{cases} \quad (\text{A1})$$

where $F(\alpha_x) = J_m(a_r r) / [\alpha_r J'_m(a_r a) - ik\beta J_m(a_r a)]$, $\sum \text{Re}^-(\alpha_x^{mn})$ is the sum of the residues of $F(\alpha_x) e^{-i\alpha_x x}$ at all its poles on and below the real axis, and $\sum \text{Re}^+(\alpha_x^{mn})$ is the sum of the residues of $F(\alpha_x) e^{-i\alpha_x x}$ at all its poles above the real axis.

The residue can be calculated by

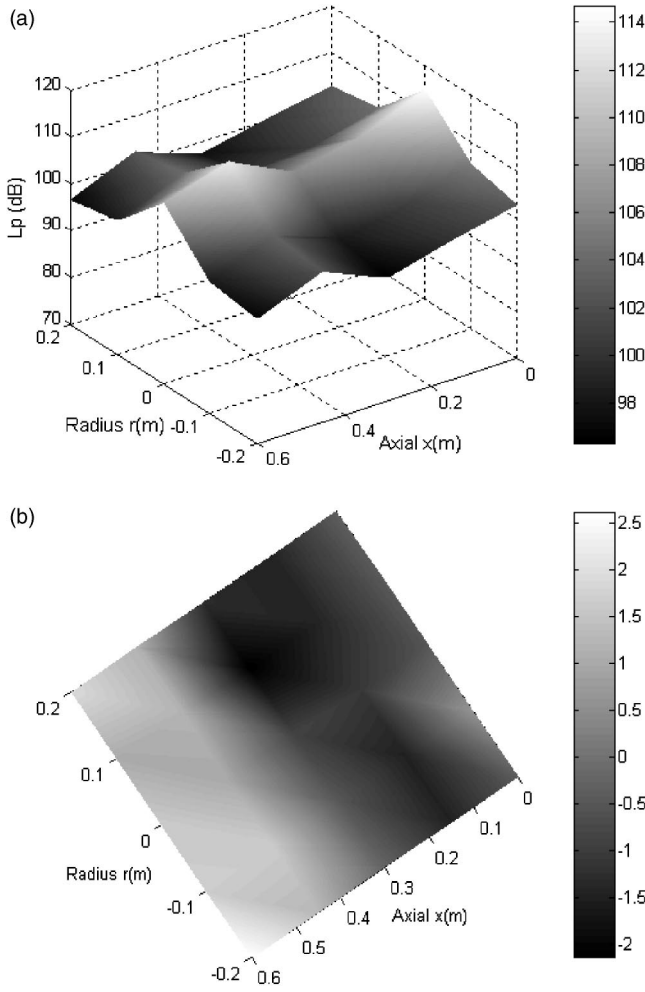


FIG. 12. The sound field (L_p) at 1850 Hz ($\beta=0.1$). (a) Calculated by the analytical model. (b) The difference of results between the analytical model and the BEM model.

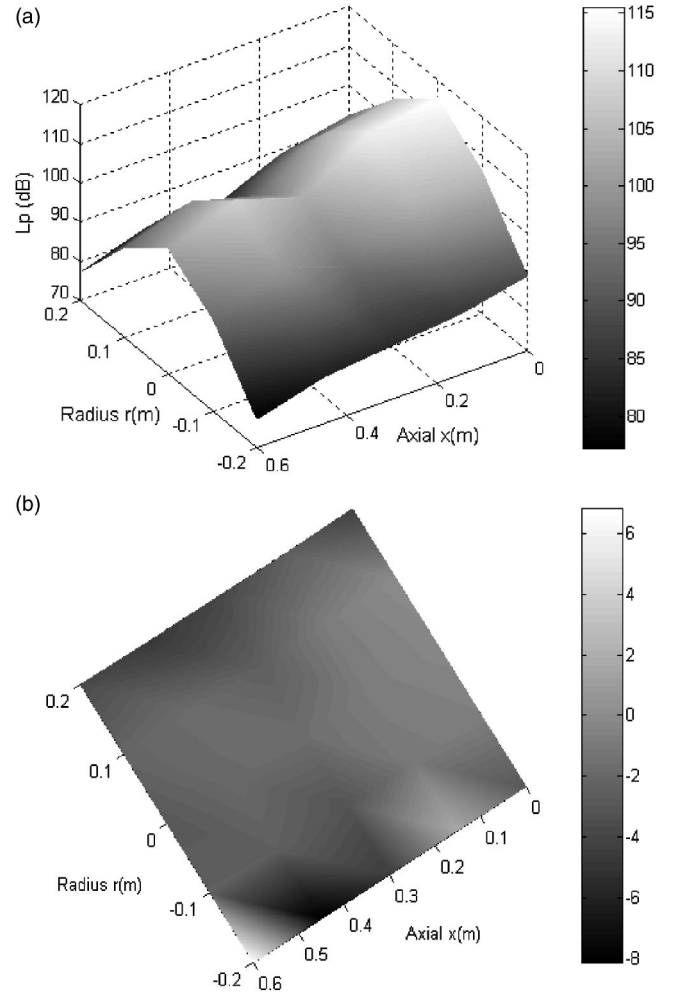


FIG. 13. The sound field (L_p) at 2400 Hz ($\beta=0.1$). (a) Calculated by the analytical model. (b) The difference of results between the analytical model and the BEM model.

$$\operatorname{Re}(\alpha_x^{mn}) = \frac{J_m(\alpha_r r)}{(d/d\alpha_x)(\alpha_r J'_m(\alpha_r a) - ik\beta J_m(\alpha_r a))|_{\alpha_x^{mn}}} = \frac{(\alpha_r^{mn})^2}{\alpha_x^{mn}} \frac{J_m(\alpha_r^{mn} r)}{(\alpha_r^{mn})^2 a J''_m(\alpha_r^{mn} a) + ik\beta J_m(\alpha_r^{mn} a)(1 - ika\beta)}. \quad (\text{A2})$$

APPENDIX B: CALCULATION OF RADIATION IMPEDANCES

The generalized radiation impedance can be written as

$$Z_{mnl} = \frac{-i}{N_{mn}N_{ml}} \int_0^\infty \frac{\tau}{\sqrt{\tau^2 - 1}} D_{mn}(\tau) D_{ml}(\tau) d\tau, \quad (\text{B1})$$

$$N_{mn} = ka \left[\frac{1}{2} \frac{((\alpha_r^{mn} a)^2 - m^2) J_m^2(\alpha_r^{mn} a)}{(\alpha_r^{mn} a)^2} + J_m'^2(\alpha_r^{mn} a) \right]^{1/2}, \quad (\text{B2})$$

$$D_{mn}(\tau) = \frac{k^2 a}{(\alpha_r^{mn})^2 - \tau^2 k^2} [k\tau J_m(\alpha_r^{mn} a) J'_m(\tau ka) - k_r^{mn} J'_m(\alpha_r^{mn} a) J_m(\tau ka)]. \quad (\text{B3})$$

Equation (B1) shows the incident modes can be coupled with other reflection modes. Equation (B1) can be simplified as

$$Z_{mnl} = \frac{1}{W} \int_0^\infty \frac{\tau^3 J_m'^2(\tau ka) - 2i\beta\tau^2 J_m(\tau ka) J'_m(\tau ka) - \beta^2 \tau J_m^2(\tau ka)}{\sqrt{1 - \tau^2} (\tau^2 - (\alpha_r^{mn})^2/k^2) (\tau^2 - (\alpha_r^{ml})^2/k^2)} d\tau \quad (\text{B4})$$

with

$$W = \frac{1}{2} \left[\left(1 - \left(\frac{\beta k}{\alpha_r^{mn}} \right)^2 - \frac{m^2}{(\alpha_r^{mn} a)^2} \right)^{1/2} \left[1 - \left(\frac{\beta k}{\alpha_r^{ml}} \right)^2 - \frac{m^2}{(\alpha_r^{ml} a)^2} \right]^{1/2} \right]. \quad (\text{B5})$$

The integral can be split into two parts over the range (0,1), and (1, ∞) and with the changes of variable, $\tau = \sin \phi$ and $\tau = \cos \xi$, in those respective ranges, the impedance equation becomes

$$Z_{mn\nu} = \frac{1}{W} \int_0^{\pi/2} \frac{\sin^3 \phi J_m'^2(ka \sin \phi) - 2i\beta \sin^2 \phi J_m(ka \sin \phi) J_m'(ka \sin \phi) - \beta^2 \sin \phi J_m^2(ka \sin \phi)}{(\sin^2 \phi - (\alpha_r^{mn})^2/k^2)(\sin^2 \phi - (\alpha_r^{m\nu})^2/k^2)} d\phi \\ + \frac{i}{W} \int_0^\infty \frac{\cosh^3 \xi J_m'^2(ka \cosh \xi) - 2i\beta \cosh^2 \xi J_m(ka \cosh \xi) J_m'(ka \cosh \xi) - \beta^2 \cosh \xi J_m^2(ka \cosh \xi)}{(\cosh^2 \phi - (\alpha_r^{mn})^2/k^2)(\cosh^2 \phi - (\alpha_r^{m\nu})^2/k^2)} d\xi. \quad (\text{B6})$$

APPENDIX C: THE MATRIX REPRESENTATION FOR EQ. (28)

For a fixed order, m , Eq. (28) can be expressed by matrices as

$$\mathbf{g}_m \mathbf{E}_m + \mathbf{E}_m' \mathbf{b}_m = \mathbf{Z}_m (\mathbf{K}_m \mathbf{E}_m'' \mathbf{b}_m - \mathbf{K}_m \mathbf{g}_m) \quad \text{for } m=0, \pm 1, \\ \pm 2, \dots, \quad (\text{C1})$$

where

$$\mathbf{g}_m = [G_{m1} \quad G_{m2} \quad G_{m\nu}]^T, \quad (\text{C2})$$

$$\mathbf{b}_m = [B_{m1} \quad B_{m2} \quad B_{m\nu}]^T. \quad (\text{C3})$$

\mathbf{E}_m , \mathbf{E}_m' , and \mathbf{E}_m'' are diagonal matrices,

$$\mathbf{E}_m = \begin{bmatrix} e^{-ia_x^{m1}L} & 0 & 0 \\ 0 & e^{-ia_x^{m2}L} & 0 \\ 0 & 0 & e^{-ia_x^{m\nu}L} \end{bmatrix}, \quad (\text{C4})$$

\mathbf{E}_m'

$$= \begin{bmatrix} e^{-ia_x^{m1}L} + e^{ia_x^{m1}L} & 0 & 0 \\ 0 & e^{-ia_x^{m2}L} + e^{ia_x^{m2}L} & 0 \\ 0 & 0 & e^{-ia_x^{m\nu}L} + e^{ia_x^{m\nu}L} \end{bmatrix}, \quad (\text{C5})$$

\mathbf{E}_m''

$$= \begin{bmatrix} e^{ia_x^{m1}L} - e^{-ia_x^{m1}L} & 0 & 0 \\ 0 & e^{ia_x^{m2}L} - e^{-ia_x^{m2}L} & 0 \\ 0 & 0 & e^{ia_x^{m\nu}L} - e^{-ia_x^{m\nu}L} \end{bmatrix}. \quad (\text{C6})$$

The general impedance matrix is given by

$$\mathbf{Z}_m = \begin{bmatrix} Z_{m11} & Z_{m12} & Z_{m1\nu} \\ Z_{m21} & Z_{m22} & Z_{m2\nu} \\ Z_{m\mu 1} & Z_{m\mu 2} & Z_{m\mu\nu} \end{bmatrix}. \quad (\text{C7})$$

\mathbf{K}_m is a diagonal matrix, which is composed of the ratio of the axial wave number of the acoustic modes to the wave number in the free field. It can be written as

$$\mathbf{K}_m = \begin{bmatrix} \alpha_x^{m1}/k & 0 & 0 \\ 0 & \alpha_x^{m2}/k & 0 \\ 0 & 0 & \alpha_x^{m\nu}/k \end{bmatrix}. \quad (\text{C8})$$

- ¹D. L. Price, J. P. D. Wilde, A. M. Papadaki, J. S. Curran, and R. I. Kitney, "Investigation of acoustic noise on 15 MRI scanners from 0.2 T to 3 T," *J. Magn. Reson. Imaging* **13**, 293–299 (2001).
- ²M. Mcjurry, R. W. Stewart, D. Crawford, and E. Toma, "The use of active noise control (ANC) to reduce acoustic noise generated during MRI scanning: Some initial results," *Magn. Reson. Imaging* **15**, 319–322 (1997).
- ³C. K. Chen, T. D. Chueh, and J. H. Chen, "Active cancellation system of acoustic noise in MR imaging," *IEEE Trans. Biomed. Eng.* **46**, 186–191 (1999).
- ⁴C. K. Mechefske and R. Geris, "Active noise control for use inside a magnetic resonance imaging machine," Ninth International Congress on Sound and Vibration (2002).
- ⁵P. Mansfield, P. Glover, and R. Bowtel, "Active acoustic screening: design principles for quiet gradient coils in MRI," *Meas. Sci. Technol.* **5**, 1021–1025 (1994).
- ⁶T. Yoshida, H. Takamori, and A. Katsunuma, "Excelart™ MRI system with revolutionary Pianssimo™ noise-reduction technology," *Med. Rev.* **71**, 1–4 (2001).
- ⁷P. Kessels, "Engineering toolbox for structural acoustic design—Applied to MRI-scanners," Ph.D. thesis, Eindhoven University of Technology, 2001.
- ⁸A. Kuijpers, S. W. Rienstra, G. Verbeek, and J. W. Verheij, "The acoustic radiation of baffled finite ducts with vibrating walls," *J. Sound Vib.* **216**, 461–493 (1998).
- ⁹P. E. Doak, "Excitation, transmission and radiation of sound from a source in a hard-walled duct of finite length, I: the effects of duct cross-section geometry and source distribution space-time pattern," *J. Sound Vib.* **31**, 1–72 (1973).
- ¹⁰P. E. Doak, "Excitation, transmission and radiation of sound from a source in a hard-walled duct of finite length, II: the effects of duct length," *J. Sound Vib.* **31**, 137–174 (1973).
- ¹¹W. E. Zorumski, "Generalized radiation impedances and reflection coefficients of circular and annular ducts," *J. Acoust. Soc. Am.* **54**, 1667–1673 (1973).
- ¹²K. S. Wang and T. C. Tzeng, "Propagation and radiation of sound in a finite length duct," *J. Sound Vib.* **93**, 57–79 (1984).
- ¹³C. R. Fuller, "Propagation and radiation of sound from flanged circular ducts with circumferentially varying wall admittances, I: Semi-infinite ducts," *J. Sound Vib.* **93**, 321–340 (1984).
- ¹⁴C. R. Fuller, "Propagation and radiation of sound from flanged circular ducts with circumferentially varying wall admittances, II: Finite ducts with sources," *J. Sound Vib.* **93**, 341–351 (1984).
- ¹⁵Y. J. Kang and I. H. Jung, "Sound propagation in circular ducts lined with noise control foams," *J. Sound Vib.* **239**, 255–273 (2001).
- ¹⁶E. G. Williams, "On Green functions for a cylindrical cavity," *J. Acoust. Soc. Am.* **102**, 3300–3307 (1997).
- ¹⁷V. Jayachanran, S. M. Hirsch, and J. Q. Sun, "On the numerical modeling of interior sound fields by the modal function expansion approach," *J. Sound Vib.* **181**, 765–780 (1993).
- ¹⁸P. M. Morse and K. Uno Ingard, *Theoretical Acoustics* (Princeton U.P., Princeton, NJ, 1986).

Acoustical measurement of the shear modulus for thin porous layers

Jean F. Allard and Michel Henry^{a)}

Laboratoire d'Acoustique de l'Université du Maine, UMR CNRS 6613, Avenue Olivier Messiaen, 72085 Le Mans Cedex, France

Laurens Boeckx, Philippe Leclaire, and Walter Lauriks

Laboratorium voor Akoestiek en Thermische Fysica, Departement Natuurkunde, Katholieke Universiteit Leuven, Celestijnenlaan 200 D, B-3001 Heverlee, Belgium

(Received 20 July 2004; revised 18 January 2005; accepted 20 January 2005)

Simulations performed with the Biot theory show that for thin porous layers, a shear mode of the structure can be induced by a point-source in air located close to the layer. The simulations show that this mode is present around frequencies where the quarter wavelength of the shear Biot wave is equal to the thickness of the samples and show that it can be acoustically detected from the fast variations with frequency of the location of a pole of the reflection coefficient close to grazing incidence. The mode has been detected with this method for two reticulated plastic foams. For one of the foams studied, the velocity and the damping of the Rayleigh wave have been measured on a thicker layer of the same medium at higher frequencies, giving a real part of the shear modulus close to the one obtained from the measured location of the pole. The strong coupling of the shear mode with the acoustic field in air allows the measurement of the shear modulus without mechanical excitation. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1868392]

PACS numbers: 43.20.Ef, 43.20.Gp, 43.20.Jr [RR]

Pages: 1737–1743

I. INTRODUCTION

Acoustical excitations from a sound source in air generally do not create noticeable displacements of the porous frame for usual sound-absorbing materials. As a consequence, the measurement of the rigidity coefficients of the porous frames is performed with mechanical excitations.^{1–7} A contact surface with a static stress is generally present and the material, often strongly nonlinearly reacting, can be modified by the experimental procedure. This drawback is avoided when the Rayleigh wave velocity is measured,⁸ but this measurement can only be performed on sufficiently thick layers; the penetration of the Rayleigh wave inside the material is of several shear wavelengths. Free field measurements of the surface impedance can be used, with the Biot theory,^{9,10} to evaluate the velocity of the frame-borne compressional wave.¹¹ At normal incidence, a mode related to this wave can be excited. This mode is with a good approximation the quarter wavelength resonance of the compressional wave in the frame, and is present at frequencies where the wavelength is close to four times the thickness l of the layer. The surface impedance at normal incidence for a layer glued to a rigid impervious backing presents around these frequencies fast variations if the loss angle of the frame is sufficiently small, but this is not the case for most of the porous sound absorbing media. The excitation of shear modes related to the Biot shear wave would need less energy at equal loss angles because the shear modulus is smaller than the modulus related to frame compressional waves. These modes cannot be excited by plane waves at normal

incidence but at large angles of incidence θ . At $\theta = \pi/2$, the surface impedance $Z_s(\pi/2)$ can be measured with the Tamura method,¹² but areas as large as 10 m² are needed, over which the porous layer must be carefully glued to a rigid impervious backing, and the experimental procedure is complicated. It has been shown recently that for thin porous layers with a motionless frame (the thickness l of a thin layer verifies $|\delta|l \ll 1$, where δ is the wave number in the air saturating the frame), a pole of the reflection coefficient exists at a complex angle of incidence θ_p close to $\pi/2$, and θ_p can be evaluated easily from measurements of the monopole pressure field close to grazing incidence.^{13–15} More precisely, the reflection coefficient V is given by

$$V(\cos \theta) = \frac{Z_s(\cos \theta) - Z/\cos \theta}{Z_s(\cos \theta) + Z/\cos \theta}, \quad (1)$$

and the angle θ_p is a solution of the following equation,

$$\cos \theta_p = -\frac{Z}{Z_s(\cos \theta_p)}, \quad (2)$$

where Z is the characteristic impedance of air, and the surface impedance Z_s is given by

$$Z_s(\theta) = \frac{iZ_1}{\phi \cos \theta_1} \cot(\delta l \cos \theta_1). \quad (3)$$

In this equation, Z_1 is the characteristic impedance in the fluid equivalent to the air saturating the porous medium, ϕ is the porosity, and θ_1 is the refraction angle given by $\sin \theta_1 = (\sin \theta)k/\delta$, k being the wave number in the free air. For thin layers θ_p is close to $\pi/2$ because $Z/|Z_s(\theta_p)| \ll 1$, and $\cos \theta_p$ is close to $-Z/Z_s(\pi/2)$. In Fig. 1, a sketch of the experimental set for the measurement method is presented. The monopole

^{a)} Author to whom correspondence should be addressed. Electronic mail: mihenry@univ-lemans.fr

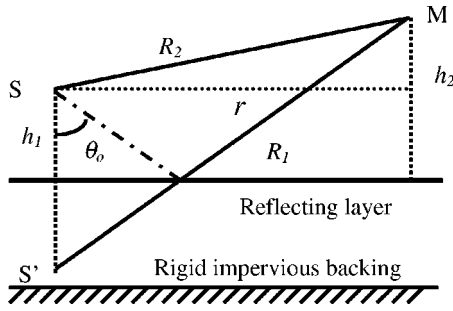


FIG. 1. Symbolic sketch of the experiment, the monopole source S, and the receiver at M above the porous layer.

source S and the receiver at M are located at small distances, respectively h_1 and h_2 , from the porous layer. The distance R_2 from the source to the receiver, R_1 from the image of the source to the receiver, and the radial distance r are much larger than $h_1 + h_2$, and the angle of specular incidence θ_0 is close to $\pi/2$. (The symbols θ_0 and θ are used for the angle of specular incidence and for the angle of incidence of plane waves, respectively.) The method is based on the use of the following approximation for the monopole reflected field p_r over thin porous layers with a motionless frame, valid under the condition θ_0 and θ_p close to $\pi/2$:

$$p_r = \frac{\exp(ikR_1)}{R_1} \left[1 - \sqrt{2\pi k R_1} \exp\left(\frac{3\pi i}{4}\right) \times \cos \theta_p \exp(-w^2) \operatorname{erfc}(-iw) \right], \quad (4)$$

$$w = \sqrt{2kR_1} \exp\left(\frac{\pi i}{4}\right) \sin \frac{\theta_p - \theta_0}{2}. \quad (5)$$

Under the condition θ_0 and θ_p close to $\pi/2$, w can be rewritten

$$w = \sqrt{\frac{1}{2}kR_1} \frac{\pi i}{4} \exp(\cos \theta_0 - \cos \theta_p). \quad (6)$$

The possible variations of $Z_s(\theta_p)$ around frequencies where the shear mode is excited are much easier to measure with the new method than $Z_s(\pi/2)$ with the Tamura method. The Biot theory is used in Sec. II to describe the shear mode and to predict the surface impedance and $\cos \theta_p$. The validity of Eqs. (4)–(6) in the context of the Biot theory is discussed in Sec. III. Measurements of $\cos \theta_p$ leading to an evaluation of the shear modulus for two porous foams are presented in Sec. IV.

II. DESCRIPTION OF THE SHEAR MODE

The Biot theory, with the formalism developed in Ref. 16, has been used to predict the surface impedance with the model by Johnson *et al.*¹⁷ for the viscous and inertial interaction, and the model by Lafarge¹⁸ for the incompressibility of air. The stress-strain relations in the Biot theory are

$$\sigma_{ij}^s = [(P - 2N)\theta^s + Q\theta^f] \delta_{ij} + 2Ne_{ij}^s, \quad (7)$$

$$\sigma_{ij}^f = Q\theta^s + R\theta^f, \quad (8)$$

where θ^f and θ^s are the dilatation of the air and of the frame, respectively, e_{ij}^s are the strain components of the frame, and σ_{ij}^f and σ_{ij}^s are the stress components of the air and of the frame, respectively. The stress components are related to forces per unit area of porous medium; then $\sigma_{ij}^f = -\phi p \delta_{ij}$, where p is the pressure in the air that saturates the porous medium. The Biot elasticity coefficients P , Q , and R , with the simplifications suggested in Ref. 16 (see Eqs. 6.21–6.28 of Ref. 16), are given by

$$P = \frac{4}{3}N + K_b + \frac{(1-\phi)^2}{\phi} K_f, \quad (9)$$

$$Q = K_f(1-\phi), \quad (10)$$

$$R = \phi K_f, \quad (11)$$

$$K_b = \frac{2}{3}N(1+\nu)/(1-2\nu), \quad (12)$$

where K_b is the bulk modulus of the frame, N is the shear modulus, and ν is the Poisson ratio. The incompressibility K_f of the air saturating the porous frame is given in the present work by

$$K_f = \frac{K_a}{\beta}, \quad (13)$$

where K_a is the adiabatic incompressibility of air and β is given by Eq. (30) of Ref. 18:

$$\beta = \gamma - (\gamma - 1) \left[1 + \frac{1}{-i\tilde{\omega}'} \left(1 - \frac{M'}{2} i\tilde{\omega}' \right)^{1/2} \right]^{-1}, \quad (14)$$

$$\tilde{\omega}' = \frac{\rho_0 \operatorname{Pr} \omega k_0'}{\eta \phi}, \quad (15)$$

$$M' = \frac{8k_0'}{\phi \Lambda'^2}. \quad (16)$$

In these equations, η is the viscosity, ρ_0 is the density of air, γ is the ratio of the specific heats, Pr is the Prandtl number, k_0' is the thermal permeability, and Λ' is the thermal characteristic dimension of the porous frame.

The equations of motion governing the displacement \mathbf{u}^s of the frame and \mathbf{u}^f of the saturating fluid are

$$-\omega^2(\tilde{\rho}_{11}\mathbf{u}^s + \tilde{\rho}_{12}\mathbf{u}^f) = (P - N)\nabla\nabla\cdot\mathbf{u}^s + N\nabla^2\mathbf{u}^s + Q\nabla\nabla\cdot\mathbf{u}^f, \quad (17)$$

$$-\omega^2(\tilde{\rho}_{22}\mathbf{u}^f + \tilde{\rho}_{12}\mathbf{u}^s) = R\nabla\nabla\cdot\mathbf{u}^f + Q\nabla\nabla\cdot\mathbf{u}^s. \quad (18)$$

The renormalized densities $\tilde{\rho}_{11}$, $\tilde{\rho}_{22}$, and $\tilde{\rho}_{12}$ are given by (see Eqs. 6.58 of Ref. 16)

$$\tilde{\rho}_{11} = \rho_1 + \rho_a - ib, \quad (19)$$

$$\tilde{\rho}_{12} = -\rho_a + ib, \quad (20)$$

$$\tilde{\rho}_{22} = \phi\rho_0 + \rho_a - ib. \quad (21)$$

The coefficient b related to the viscous and the inertial interaction is given in the present work by

$$b = -\frac{\sigma}{\omega} \phi^2 G_j(\omega), \quad (22)$$

$$G_j(\omega) = \left(1 - \frac{4i\alpha_\infty^2 \eta \rho_0 \omega}{\sigma^2 \Lambda^2 \phi^2} \right)^{1/2} \quad (23)$$

(see Eq. 5–64 of Ref. 16 with a complex conjugation due to the opposite time dependence), where α_∞ is the tortuosity, σ is the flow resistivity, and Λ is the characteristic viscous dimension. The added density ρ_a is given by

$$\rho_a = \rho_0 \phi (\alpha_\infty - 1). \quad (24)$$

The wave numbers k_1 and k_2 for both Biot compressional waves, and k_3 for the Biot shear wave, are given by

$$k_{1,2}^2 = \frac{\omega^2}{2(PR - Q^2)} [P\tilde{\rho}_{22} + R\tilde{\rho}_{11} - 2Q\tilde{\rho}_{12} \pm \sqrt{\Delta}], \quad (25)$$

$$k_3^2 = \frac{\omega^2}{N} \left[\frac{\tilde{\rho}_{11}\tilde{\rho}_{22} - \tilde{\rho}_{12}^2}{\tilde{\rho}_{22}} \right], \quad (26)$$

$$\Delta = (P\tilde{\rho}_{22} + R\tilde{\rho}_{11} - 2Q\tilde{\rho}_{12})^2 - 4(PR - Q^2)(\tilde{\rho}_{11}\tilde{\rho}_{22} - \tilde{\rho}_{12}^2). \quad (27)$$

Both compressional waves can be related to a velocity potential φ for the frame and $\varphi_f = \mu \varphi$ for the saturating fluid. The related displacements \mathbf{u}^s and \mathbf{u}^f are given by

$$\mathbf{u}^s = \frac{i}{\omega} \nabla \varphi, \quad (28)$$

$$\mathbf{u}^f = \frac{i\mu}{\omega} \nabla \varphi. \quad (29)$$

The coefficients μ are given by

$$\mu_{1,2} = \frac{Pk_{1,2}^2 - \omega^2 \tilde{\rho}_{11}}{\omega^2 \tilde{\rho}_{12} - Qk_{1,2}^2}. \quad (30)$$

The shear wave is related to a vector velocity potential for the frame Ψ , and a vector velocity potential for air $\Psi_a = \mu_3 \Psi$.

The related displacements are given by

$$\mathbf{u}^s = \frac{i}{\omega} \nabla \wedge \Psi, \quad (31)$$

$$\mathbf{u}^f = \frac{i}{\omega} \mu_3 \nabla \wedge \Psi, \quad (32)$$

and μ_3 is given by

$$\mu_3 = -\tilde{\rho}_{12}/\tilde{\rho}_{22}. \quad (33)$$

A plane wave in air impinging upon the porous layer with an angle of incidence θ is represented in Fig. 2. The incidence plane is xoz , the axis x is parallel to the air-porous medium interface. The incident field creates in the porous layer a field which can be described by six potentials with the same horizontal wave number vector component $\xi = k \sin \theta$, φ_1^+ , φ_1^- for the first compressional wave with opposite z wave number vector components $\alpha_1 = (k_1^2 - \xi^2)^{1/2}$ and $-\alpha_1$, respectively. Similarly, two potentials φ_2^+ and φ_2^- are defined for the second compressional wave, related to the z wave number vector components $\alpha_2 = (k_2^2 - \xi^2)^{1/2}$ and $-\alpha_2$, and two potentials $\Psi^+ = \mathbf{n}\varphi_3^+$, $\Psi^- = \mathbf{n}\varphi_3^-$, related to $\alpha_3 = (k_3^2$

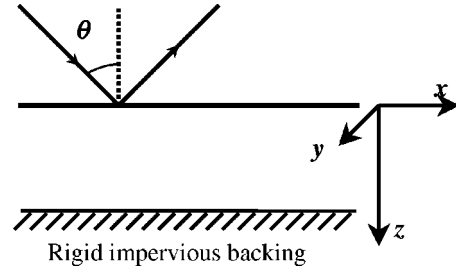


FIG. 2. A plane wave in air, incident on a porous layer. The angle of incidence is θ , the incidence plane is oxz .

$-\xi^2)^{1/2}$ and $-\alpha_3$, \mathbf{n} being the unit vector on the y axis. The scalar functions φ_j^\pm can be written

$$\varphi_j^\pm = a_j^\pm \exp(\pm i\alpha_j z + i\xi x), \quad (34)$$

where the six constant coefficients a_j^\pm can be predicted from the boundary conditions.

The porous layer is glued to the rigid impervious backing. At the contact surface with the rigid impervious backing the boundary conditions are $u_z^s = u_x^s = u_z^f = 0$, which can be rewritten

$$\begin{aligned} \alpha_1 [a_1^+ \exp(i\alpha_1 l) - a_1^- \exp(-i\alpha_1 l)] + \alpha_2 [a_2^+ \exp(i\alpha_2 l) \\ - a_2^- \exp(-i\alpha_2 l)] + \xi [a_3^+ \exp(i\xi l) + a_3^- \exp(-i\xi l)] \\ = 0, \end{aligned} \quad (35)$$

$$\begin{aligned} \xi [a_1^+ \exp(i\alpha_1 l) + a_1^- \exp(-i\alpha_1 l)] + \xi [a_2^+ \exp(i\alpha_2 l) \\ + a_2^- \exp(-i\alpha_2 l)] - \alpha_3 [a_3^+ \exp(i\xi l) - a_3^- \exp(-i\xi l)] \\ = 0, \end{aligned} \quad (36)$$

$$\begin{aligned} \alpha_1 \mu_1 [a_1^+ \exp(i\alpha_1 l) - a_1^- \exp(-i\alpha_1 l)] \\ + \alpha_2 \mu_2 [a_2^+ \exp(i\alpha_2 l) - a_2^- \exp(-i\alpha_2 l)] \\ + \xi \mu_3 [a_3^+ \exp(i\xi l) + a_3^- \exp(-i\xi l)] = 0. \end{aligned} \quad (37)$$

Let p and U_z be the pressure and the z component of the air velocity in the free air at the contact surface with the porous layer. At this interface, the boundary conditions can be written¹⁹

$$\phi u_z^f + (1 - \phi) u_z^s = U_z, \quad (38)$$

$$\sigma_{zz}^f = -\phi p, \quad (39)$$

$$\sigma_{zz}^s = -(1 - \phi)p, \quad (40)$$

$$\sigma_{xz}^s = 0. \quad (41)$$

A preliminary use of Eqs. (35)–(37) provides the reflection coefficients at a rigid and impervious boundary for the three Biot waves and first information concerning the shear modes. For $a_1^+ = 1$, $a_2^+ = a_3^+ = 0$, i.e., a compressional wave φ_1^+ of amplitude unity at the air-porous layer boundary, the expressions for the a_i^- , $i = 1, 3$, solutions of Eqs. (35)–(37) and denoted as $r_{1,i}$ are given in the first line of Table I. In the same way, the second line gives the expressions $r_{2,i}$, $i = 1, 3$, of the a_i^- related to $a_2^+ = 1$, $a_1^+ = a_3^+ = 0$, and the third line gives the expressions $r_{3,i}$, $i = 1, 3$, of the a_i^- related to a_3^+

TABLE I. The coefficients $r_{i,j}$ expressed with $A = \alpha_1 \alpha_2 (\mu_1 - \mu_2)$, $B = \alpha_1 \xi (\mu_1 - \mu_3)$, $C = \alpha_2 \xi (\mu_2 - \mu_3)$, $D = \alpha_3 A + \xi (B - C)$, and $T_{i,j} = \exp(-i(\alpha_i + \alpha_j)l)$.

$r_{1,1}DT_{1,1}$ $\alpha_3 A + \xi(B + C)$	$r_{1,2}DT_{1,2}$ $-2\xi B$	$r_{1,3}DT_{1,3}$ $-2\xi A$
$r_{2,1}DT_{2,1}$ $2\xi C$	$r_{2,2}DT_{2,2}$ $\alpha_3 A - \xi(B + C)$	$r_{2,3}DT_{2,3}$ $-2\xi A$
$r_{3,1}DT_{3,1}$ $-2\alpha_3 C$	$r_{3,2}DT_{3,2}$ $2\alpha_3 B$	$r_{3,3}DT_{3,3}$ $\alpha_3 A - \xi(B - C)$

$= 1$, $a_1^+ = a_2^+ = 0$. The factor $\exp(i\xi x)$ has been removed from the different amplitudes. At the air-porous layer interface, each incident wave in the layer will be associated with its reflected field, so that Eqs. (35)–(37) will be automatically satisfied. Equations (38)–(41) can be rewritten

$$i\alpha_1(\phi\mu_1 + (1 - \phi))[a_1^+(1 - r_{1,1}) - a_2^+r_{2,1} - a_3^+r_{3,1}] + i\alpha_2(\phi\mu_2 + (1 - \phi))[-a_1^+r_{1,2} + a_2^+(1 - r_{2,2}) - a_3^+r_{3,2}] + i\xi(\phi\mu_3 + (1 - \phi))[a_1^+r_{1,3} + a_2^+r_{2,3} + a_3^+(1 + r_{3,3})] = -i\omega U_z, \quad (42)$$

$$\frac{-i}{\omega}\{(Q + R\mu_1)k_1^2[a_1^+(1 + r_{1,1}) + a_2^+r_{2,1} + a_3^+r_{3,1}] + (Q + R\mu_2)k_2^2[a_1^+r_{1,2} + a_2^+(1 + r_{2,2}) + a_3^+r_{3,2}]\} = -\phi p, \quad (43)$$

$$\frac{-i}{\omega}\{(P + Q\mu_1)k_1^2 - 2N\xi^2[a_1^+(1 + r_{1,1}) + a_2^+r_{2,1} + a_3^+r_{3,1}] + [(P + Q\mu_2)k_2^2 - 2N\xi^2][a_1^+r_{1,2} + a_2^+(1 + r_{2,2}) + a_3^+r_{3,2}] + 2N\xi\alpha_3[-a_1^+r_{1,3} - a_2^+r_{2,3} + a_3^+(1 - r_{3,3})]\} = -(1 - \phi)p, \quad (44)$$

$$2\xi\alpha_1[a_1^+(1 - r_{1,1}) - a_2^+r_{2,1} - a_3^+r_{3,1}] + 2\xi\alpha_2[-a_1^+r_{1,2} + a_2^+(1 - r_{2,2}) - a_3^+r_{3,2}] + (\xi^2 - \alpha_3^2) \times [a_1^+r_{1,3} + a_2^+r_{2,3} + a_3^+(1 + r_{3,3})] = 0. \quad (45)$$

The three Eqs. (43)–(45) provide the ratios a_1^+/p , a_2^+/p , and a_3^+/p . Reporting these ratios in Eq. (42) provides the surface impedance $Z_s(\theta) = p/U_z$. The predicted $\cos\theta_p$ is obtained from the series $(\cos\theta_p)_i$ defined by the following iterative equation

$$(\cos\theta_p)_{i+1} = 0.8(\cos\theta_p)_i - 0.2 \frac{Z}{Z_s((\theta_p)_i)}, \quad (46)$$

with $(\theta_p)_1 = \pi/2$. A general description of the shear mode with Eqs. (42)–(45) is beyond the aim of the present work. The main properties of the mode are shown in the context of a simulation with a porous foam.

The porous material used for the simulations is a typical foam with a thickness $l = 1$ cm. The parameters that describe the foam are given in Table II (yellow foam), except the imaginary part of the shear modulus N , and the thickness l . The layer is set on a rigid impervious backing and the frame

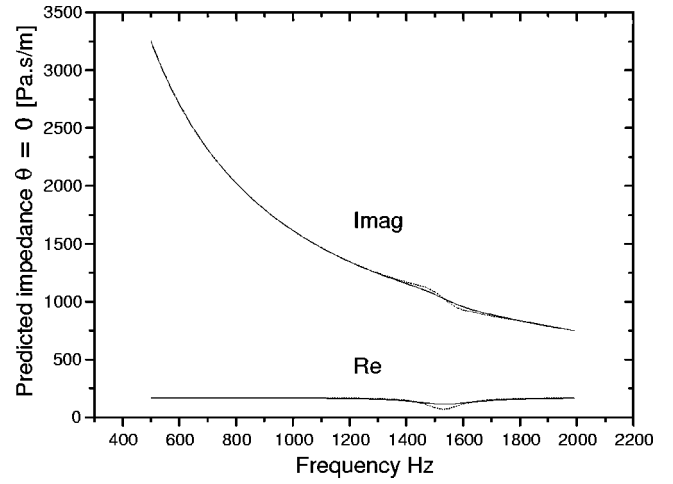


FIG. 3. The surface impedance at normal incidence. Same parameters as in the second column of Table I (yellow foam) except thickness $l = 1$ cm and $\text{Im } N = -6.2$ kPa: —, $\text{Im } N = -1.55$ kPa: ---.

is motionless at the contact surface. The surface impedances at normal incidence $Z_s(0)$ and $\cos(\theta_p)$ are represented in Figs. 3 and 4 for two loss angles, 0.1 and 0.025, respectively. The Poisson ratio ν is real, because previous measurements^{7,20} give $\text{Im } \nu = 0$, and $\text{Re } \nu = 0.3$. This choice for $\text{Re } \nu$ is arbitrary; coefficients ν previously measured for different foams range from 0.1 to 0.5, but it has no significant effect on the shear mode. There is only a small shift, around 10 Hz, of the location of the peak related to the shear mode in Fig. 4 when ν varies from 0 to 0.5. The shear mode is located around 813 Hz, where the quarter shear Biot wavelength is equal to the thickness of the layer, and the compressional mode is located around 1500 Hz. At normal incidence, $\xi = 0$, and $r_{3,1} = r_{3,2} = r_{1,3} = r_{2,3} = 0$, $r_{3,3} = 1$. Then Eq. (45) gives $a_3^+ = 0$ and the Biot shear wave is not present. Let ξ_p be the horizontal wave number component related to the pole, $\xi_p = k \sin\theta_p$. When $\cos\theta_p$ is small, $\sin\theta_p$ is close to 1, and ξ_p is close to k . All the coefficients $r_{i,j}$ are different from 0, and the shear wave is present and gives a contribution to U_z and p via u_z^s , u_z^f , σ_{zz}^s , and σ_{zz}^f . This contribution and a_3^+ reach a maximum when the coefficient $T_{3,3} = \exp(-2i\alpha_3 l)$ of

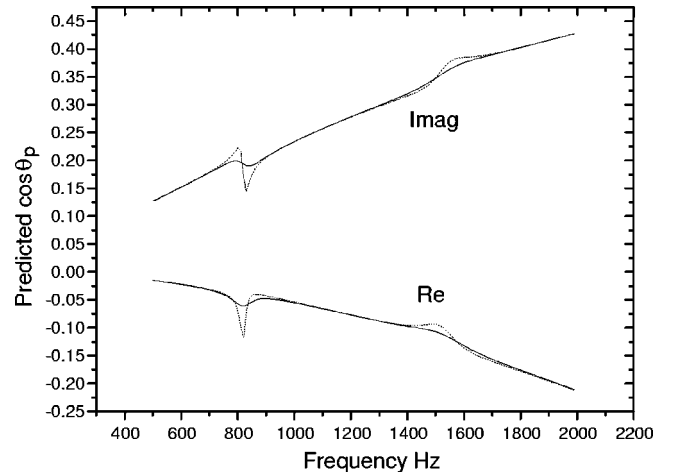


FIG. 4. Predicted $\cos\theta_p$ as a function of frequency for the same material as in Fig. 3. $\text{Im } N = -6.2$ kPa: —, $\text{Im } N = -1.55$ kPa: ---.

Table I is close to -1 , for $\text{Re } \alpha_3 l$ close to $\pi/2$. The shear mode does not exist if the shear wave is faster than the sound in air, because in this case $\alpha_3 = (k_3^2 - k^2)^{1/2}$ is, in a first approximation, an imaginary quantity. In the simulation (and for the usual plastic open-cell foams), the speed of the shear waves in the foam is much smaller than the sound speed in the free air, and α_3 is close to k_3 . For a loss angle equal to 0.1 ($\text{Im } N = -6200 \text{ Pa}$), the influence of the compressional mode is negligible on the impedance but the shear mode is detectable on the behavior of $\cos \theta_p$. For a loss angle equal to 0.025 ($\text{Im } N = -1550 \text{ Pa}$), the presence of the shear mode is much more obvious than the presence of the compressional mode on the pole location. The shear mode appears in a small range of frequencies where it is related to a horizontal wave number component close to k , which allows a strong coupling with the pressure field in the free air at grazing incidence. An opposite case is the one of the Rayleigh wave, which can propagate at the same frequencies on samples of larger thickness. The order of magnitude of the Rayleigh wave velocity is around 50 m/s, which corresponds to a horizontal wave number component seven times as large as k . The coupling with the field in air at grazing incidence is weak, with no noticeable effect on $\cos \theta_p$.

III. VALIDITY OF THE APPROXIMATION FOR p_r IN THE CONTEXT OF THE BIOT THEORY

The method used to measure $\cos \theta_p$ for thin porous layers with a motionless frame is based on the use of Eqs. (4)–(6). These expressions have been obtained with the reference integral method described by Brekhovskikh and Godin²¹ from the Sommerfeld representation of the reflected monopole field

$$p_r = i \int_0^\infty \frac{k \xi d\xi}{\lambda} V(\lambda) J_0(\xi k r) \exp[ik\lambda(h_1 + h_2)], \quad (47)$$

which can be rewritten

$$p_r = \frac{i}{2} \int_{-\infty}^\infty \frac{k \xi d\xi}{\lambda} V(\lambda) H_0^{(1)}(\xi k r) \exp[ik\lambda(h_1 + h_2)], \quad (48)$$

$$\lambda = (1 - \xi^2)^{1/2}, \quad \text{Im}(\lambda) \geq 0,$$

where $H_0^{(1)}$ is the Hankel function of zeroth order. A complex angle θ can be defined by $\sin \theta = \xi$, $\cos \theta = \lambda$. Using the asymptotic development of the Hankel function, $H_0^{(1)}(u) = (2/\pi u)^{1/2} \exp(iu - i\pi/4)$, Eq. (48) can be rewritten

$$p_r = \left(\frac{k}{2\pi r}\right)^{1/2} \exp\left(\frac{i\pi}{4}\right) \int_{-\infty}^\infty F(\xi) \exp[\rho f(\lambda)] d\xi, \quad (49)$$

$$F(\xi) = \sqrt{\frac{\xi}{1 - \xi^2}} V(\lambda), \quad (50)$$

$$f(\lambda) = i \cos(\theta - \theta_0), \quad (51)$$

$$\rho = kR_1. \quad (52)$$

Using the reference integral method with the notations of Ref. 21, $q = \sin \theta$, $q_p = \sin \theta_p$, and $q_0 = \sin \theta_0$, p_r can be rewritten

$$p_r = \left(\frac{k}{2\pi r}\right)^{1/2} \exp(i\pi/4) \exp[\rho f(q_0)] \times \left[a \mathcal{F}_1(1, \rho, s_p) + \left(\frac{\pi}{\rho}\right)^{1/2} \Phi_1^0(0) \right], \quad (53)$$

$$a = 1 / \left(\frac{dF^{-1}(q)}{dq} \right)_{q=q_p}, \quad (54)$$

$$\mathcal{F}_1(1, \rho, s_p) = i\pi \exp(-\rho s_p^2) \text{erfc}(-i\sqrt{\rho} s_p), \quad (55)$$

$$\Phi_1^0(0) = F(q_0) (-2/f_{q_2}''(q_0))^{1/2} + \frac{a}{s_p}, \quad (56)$$

$$s_p = \exp(i\pi/4) \sqrt{2} \sin \frac{\theta_p - \theta_0}{2}. \quad (57)$$

With $f_{q_2}''(q_0) = -i/\cos^2(\theta_0)$, the term $F(q_0) (-2/f_{q_2}''(q_0))^{1/2}$ gives a contribution p_{r1} to p_r given by

$$p_{r1} = V(\cos \theta_0) \exp(ikR_1) / R_1, \quad (58)$$

independently on the expression which defines V . The two terms with the factor a give the contribution p_{r2} given by

$$p_{r2} = a \exp(ikR_1 + i3\pi/4) (k/2r)^{1/2} (-i + w\sqrt{\pi} \exp(-w^2) \times \text{erfc}(-iw)) / w, \quad (59)$$

where w is given by Eq. (5), and the coefficient a is given by

$$a = \frac{-2 \cos \theta_p}{\sqrt{\sin \theta_p}} [1 + (\cos \theta_p / Z_s(\cos \theta_p))] \times (dZ_s/d \cos \theta)_{\theta=\theta_p}^{-1}. \quad (60)$$

The reflected pressure can be written

$$p_r = V(\cos \theta_0) \exp(ikR_1) / R_1 - \frac{\cos \theta_p}{w} \sqrt{\frac{2k}{r \sin \theta_p}} \times \frac{\exp(ikR_1 + i3\pi/4) [-i + w\sqrt{\pi} \exp(-w^2) \text{erfc}(-iw)]}{[1 + (\cos \theta_p / Z_s(\cos \theta_p)) (dZ_s/d \cos \theta)_{\theta=\theta_p}]}. \quad (61)$$

If Z_s does not depend on θ_p , this equation is identical to Eq. 1-4-10 of Ref. 21 for locally reacting surfaces

$$p_r = V(\cos \theta_0) \exp(ikR_1) / R_1 - \frac{\cos \theta_p}{w} \sqrt{\frac{2k}{r \sin \theta_p}} \exp(ikR_1 + i3\pi/4) [-i + w\sqrt{\pi} \exp(-w^2) \text{erfc}(-iw)]. \quad (62)$$

If θ_p and θ_0 are sufficiently close to $\pi/2$, ($\sin \theta_p$)^{1/2} can be removed and Eqs. (5) and (62) can be replaced by Eqs. (6) and (4). These equations can be used also for thin layers with a motionless frame or in the context of the Biot theory if

$$|(\cos \theta_p / Z_s(\cos \theta_p)) (dZ_s/d \cos \theta)_{\theta=\theta_p}| \ll 1. \quad (63)$$

An evaluation of the left side of this equation is not very simple. If the structural damping of the frame is small, the shear mode induces fast variations of Z_s with frequency, Eq. (63) is not satisfied, and Eqs. (4)–(6) cannot be used, but the loss angle for usual open-cell foams is large. Comparisons

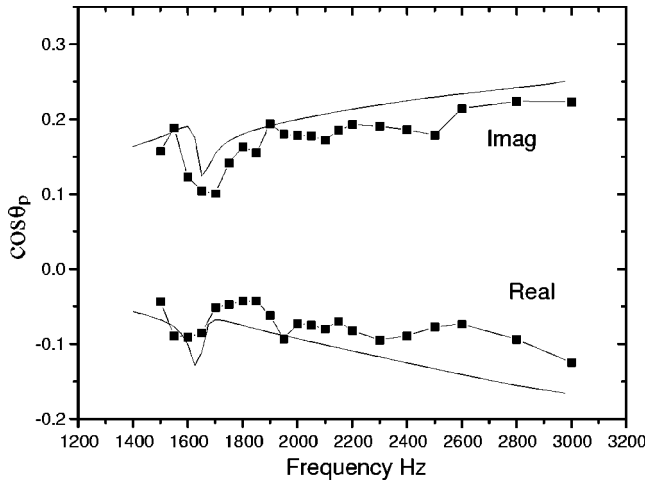


FIG. 5. $\cos \theta_p$ as a function of frequency, yellow foam; measurement \blacksquare - \blacksquare - \blacksquare , prediction —.

between estimations of the reflected pressure over thin porous layers, obtained with Eq. (47) and with Eqs. (4)–(6), have been performed around frequencies where the shear mode is present, showing that for loss angles of the frame larger than $\frac{1}{100}$, Eqs. (4) and (6) provide a good approximation for the reflected pressure for thin layers. The loss angle for usual open-cell foams is much larger than $\frac{1}{100}$.

IV. EXPERIMENTAL EVALUATION OF THE SHEAR MODULUS

In Fig. 1 a symbolic sketch of the experimental set up is presented. The monopole source S is a pipe fed by a compression driver. The receiver at M is an electret microphone. The distances z_1 from the porous layer to the source and z_2 from the layer to the microphone are $z_1 = z_2 = 3$ mm. The distance r from the source to the receiver is $r = 23$ cm. The angle of specular reflection is θ_0 verifying $\cot \theta_0 = (z_1 + z_2)/r$. The porous layers are glued to a rigid impervious backing, the area of the layers is close to 1 m^2 . The evaluation of $\cos \theta_p$ with Eq. (4) is performed from a measurement of the ratio $p_r / (\exp(ikR_1)/R_1)$ where $\exp(ikR_1)/R_1$ is the direct field at a distance R_1 from the source. The measurement of the ratio is obtained from a measurement of the pressure p_h related to the direct field, the porous surface being replaced by a rigid, impervious surface and a second measurement over the porous surface of the pressure p_t . Both measurements are performed with the same geometry, close to grazing incidence so that R_1 and R_2 are equal in a first approximation. The ratio in Eq. (4) can be evaluated from

$$p_r / (\exp(ikR_1)/R_1) = \frac{p_t - p_h/2}{p_h/2}, \quad (64)$$

and $\cos \theta_p$ is given by

$$2 \frac{p_h - p_t}{p_h} = \cos \theta_p \exp\left(\frac{3\pi i}{4}\right) \left[\sqrt{2\pi k R_1} \exp(-w^2) \times \text{erfc}(-iw) \right]. \quad (65)$$

An iterative procedure¹⁴ is used to evaluate $\cos \theta_p$ from Eq. (65). For thin layers with a source-receiver

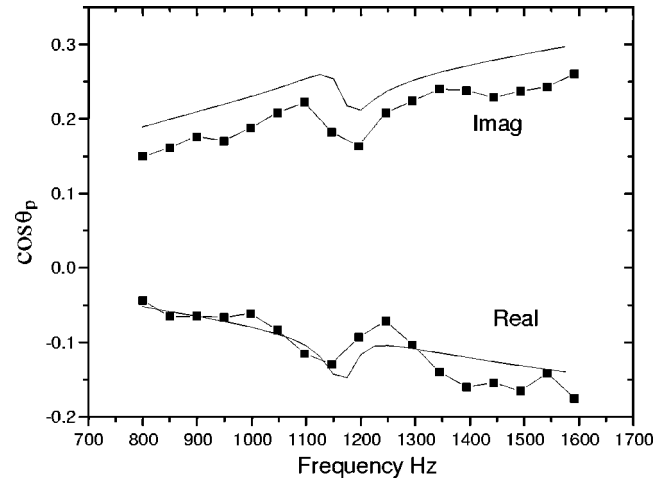


FIG. 6. $\cos \theta_p$ as a function of frequency, black foam; measurement \blacksquare - \blacksquare - \blacksquare , prediction —.

distance smaller than 1 m, w is small and the first value for $\cos \theta_p$ is obtained with $w = 0$, giving $\cos \theta_p = 2(1 - p_t/p_h) / \exp(3\pi i/4) \sqrt{2\pi k R_1}$. The derivative of $g(w) = \exp(-w^2) \text{erfc}(-iw)$ is $g'_w(w) = -2wg(w) + 2i/\sqrt{\pi}$ (see Ref. 22, Chap. 7), and the derivative $dw/d \cos \theta_p$ of $w = (1/2ikR_1)^{1/2} (\cos \theta_0 - \cos \theta_p)$ is $dw/d \cos \theta_p = -(1/2ikR_1)^{1/2}$. These two relations are used to minimize the left- and right-hand side difference of Eq. (65).

Measurements have been performed on two foams, a yellow foam of thickness $l = 0.5$ cm and a black foam (S616D) manufactured by Recticel (Plejadenaan, 15, B-1200 Brussels) of thickness $l = 1$ cm. Measurements on the yellow foam are presented in Fig. 5 and on the black foam in Fig. 6. Fast variations of $\cos \theta_p$ similar to the ones of Fig. 4 are present for both foams. A set of parameters given in Table II has been used to predict $\cos \theta_p$ in Figs. 4 and 5. The flow resistivities and the densities in Table II are measured and the shear moduli have been chosen to set the fast predicted variations at the same frequencies as the measured ones. The other parameters have been set to reasonable values. The frequencies where the fast variations occur are mainly related to densities and shear moduli; the other parameters have a very small influence on the localization of the fast variations. For the black foam, measurements of the Rayleigh wave velocity have been performed on a sample of thickness $l = 6$ cm with the method described in Ref. 8: a Rayleigh-like wave is excited with a conventional shaker in

TABLE II. Parameters for the porous materials.

Parameters	Yellow foam	Black foam
porosity ϕ	0.99	0.99
flow resistivity σ ($\text{Nm}^{-4} \text{s}$)	40 000.0	130 000.0
thermal permeability k'_0 (m^2)	10^{-9}	5×10^{-10}
viscous dimension Λ (μm)	100	60
thermal dimension Λ' (μm)	300	180
tortuosity α_∞	1.05	1.1
density ρ_s (kg/m^3)	58	59
shear modulus N (kPa)	62.0 - i 1.55	125.0 - i 6.0
Poisson ratio (ν)	0.3	0.3
thickness l (cm)	0.5	1.0

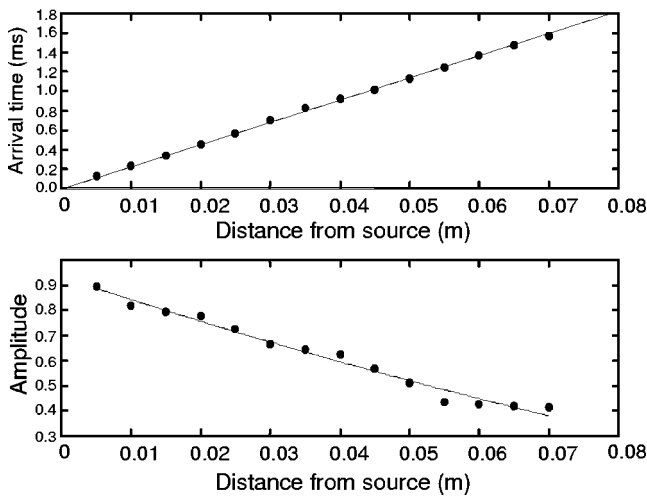


FIG. 7. Arrival time and attenuation as a function of position at 2 kHz, 25 °C. The phase velocity and attenuation are determined from a linear regression and an exponential regression on the data, respectively.

contact with the surface of the sample and the vertical displacement is measured at several distances from the source using a Polytec vibrometer. A line source has been used as an excitation and the signal was a narrow-band Gaussian burst centered around the measuring frequency. The shear modulus was extracted from the phase velocity of the Rayleigh wave and the attenuation was measured from the decay of the amplitude as a function of position. Figures 7(a) and (b) show typical results as a function of position at a center frequency of 2 kHz and ambient temperature 25 °C. Table III summarizes the results at 2 and 4 kHz under the hypothesis of a Poisson ratio $\nu=0.3$. The yellow foam is much less rigid at very low frequencies than the black foam, but it has not been possible to measure the Rayleigh wave velocity (samples of sufficient thickness were not available). The effect of the shear mode is a peak limited over a small range of frequencies. The precision of the measurement of $\text{Re } N$ is limited by the broadness of the peak. For the materials studied, the precision is about $\frac{10}{100}$ for $\text{Re } N$ and $\frac{50}{100}$ for $\text{Im } N$. The fast increase of $\text{Im } N$ with frequency can explain the difference between the measured $\text{Im } N$ around 1.2 kHz with the point source method and around 2 kHz with the Rayleigh wave. The measurements of $\text{Re } N$ by both methods give similar results.

V. CONCLUSION

The acoustic field created by a point-source close to thin layers of open-cell foams can be strongly coupled with a shear mode around frequencies where the quarter shear wavelength is equal to the thickness of the layer. The strong coupling is due to the fact that the horizontal wave-number

TABLE III. Measured wave number of the Rayleigh wave and shear modulus.

Frequency (kHz)	$\text{Re } k_R$ (m^{-1})	$\text{Im } k_R$ (m^{-1})	N (kPa)
2	283.0	12.0	135.0 - i 11.0
4	576.0	65.0	126.0 - i 29.0

component of the mode is close to the wave number in the free air. The amplitude of the mode is sufficiently large for the mode to be detected acoustically from the fast variations with frequency of the location of a pole of the reflection coefficient. The description with the Biot theory of the interaction between the shear mode and the acoustic field in air is in good agreement with the experiment. This is a new example of the adequacy of the Biot theory for the description of the acoustical properties of sound-absorbing media. This work also shows the interest of the method of localizing a pole of the reflection coefficient from short range pressure measurements close to grazing incidence.

- ¹T. Pritz, "Transfer function method for investigating the complex modulus of acoustic materials: spring-like specimen," *J. Sound Vib.* **72**, 317–341 (1980).
- ²T. Pritz, "Transfer function method for investigating the complex modulus of acoustic materials: rod-like specimen," *J. Sound Vib.* **81**, 359–376 (1982).
- ³T. Pritz, "Frequency dependence of frame dynamic characteristics of mineral and glass wool materials," *J. Sound Vib.* **106**, 161–169 (1986).
- ⁴T. Pritz, "Dynamic Young's modulus and loss factor of plastic foams for impact sound isolation," *J. Sound Vib.* **178**, 315–322 (1994).
- ⁵A. Sfaoui, "On the viscoelasticity of the polyurethane foam," *J. Acoust. Soc. Am.* **97**, 1046–1052 (1995).
- ⁶T. Pritz, "Measurement methods of complex Poisson's ratio of viscoelastic materials," *Appl. Acoust.* **60**, 279–292 (2000).
- ⁷M. Melon, M. Mariez, C. Ayrault, and S. Sahaoui, "Acoustical and mechanical characterization of anisotropic open-cell foams," *J. Acoust. Soc. Am.* **104**, 2622–2627 (1998).
- ⁸J. F. Allard, G. Jansens, G. Vermeir, and W. Lauriks, "Frame-borne surface waves in air-saturated porous media," *J. Acoust. Soc. Am.* **111**, 690–696 (2002).
- ⁹M. A. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid," *J. Acoust. Soc. Am.* **28**, 168–191 (1956).
- ¹⁰M. A. Biot and D. G. Willis, "The elastic coefficients of the theory of consolidation," *J. Appl. Mech.* **24**, 594–601 (1957).
- ¹¹J. F. Allard, C. Depollier, P. Guignouard, and P. Rebillard, "Effect of a resonance of the frame on the surface impedance of a glass-wool of high density and stiffness," *J. Acoust. Soc. Am.* **89**, 999–1001 (1991).
- ¹²M. Tamura, "Spatial Fourier transform method of measuring reflection coefficient at oblique incidence," *J. Acoust. Soc. Am.* **88**, 2259–2264 (1990).
- ¹³J. F. Allard, M. Henry, V. Garetton, J. Jansens, and W. Lauriks, "Impedance measurement around grazing incidence for nonlocally reacting thin porous layers," *J. Acoust. Soc. Am.* **113**, 1210–1215 (2003).
- ¹⁴J. F. Allard, V. Garetton, M. Henry, G. Jansens, and W. Lauriks, "Impedance evaluation from pressure measurements near grazing incidence for nonlocally reacting porous layers," *Acust. Acta Acust.* **89**, 595–603 (2003).
- ¹⁵J. F. Allard, M. Henry, and V. Garetton, "Pseudo-surface waves above thin porous layers," *J. Acoust. Soc. Am.* **116**, 1345–1347 (2004).
- ¹⁶J. F. Allard, *Propagation of Sound in Porous Media, Modelling Sound Absorbing Materials* (Elsevier, London, 1993).
- ¹⁷D. L. Johnson, J. Koplik, and R. Dashen, "Theory of dynamic permeability and tortuosity in fluid-saturated porous media," *J. Fluid Mech.* **176**, 379–402 (1987).
- ¹⁸D. Lafarge, P. Lemarinier, J. F. Allard, and V. Tarnow, "Dynamic compressibility of air in porous structures at audible frequencies," *J. Acoust. Soc. Am.* **102**, 1995–2006 (1997).
- ¹⁹H. Deresiewicz and R. Skalak, "On uniqueness in dynamic poroelasticity," *Bull. Seismol. Soc. Am.* **53**, 783–788 (1963).
- ²⁰E. Mariez, S. Sahaoui, and J. F. Allard, "Elastic constants of polyurethane foam's skeleton for Biot model," *Internoise Congress Proceedings, Budapest (1997)*, pp. 1683–1686.
- ²¹L. M. Brekhovskikh and O. A. Godin, *Acoustics of Layered Media II, Point Source and Bounded Beams*, Springer Series on Wave Phenomena (Springer, New York, 1992).
- ²²M. Abramovitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970).

Computation of scattering from clusters of spheres using the fast multipole method

Nail A. Gumerov and Ramani Duraiswami

*Perceptual Interfaces and Reality Laboratory, Institute for Advanced Computer Studies,
University of Maryland, College Park, Maryland 20742*

(Received 23 June 2004; revised 5 November 2004; accepted for publication 1 December 2004)

A T-matrix based method of solution of the multiple scattering problem was presented by the authors [J. Acoust. Soc. Am. **112**, 2688–2701 (2002)]. This method can be applied to the computation of relatively small problems, since the number of operations required grows with the number of spheres N as $O(N^3)$, and with the sixth power of the wave number. The use of iterative techniques accelerated using the fast multipole method (FMM) can accelerate this solution, as presented by Koc and Chew [J. Acoust. Soc. Am. **103**, 721–734 (1998)] originally. In this study we present a method that combines preconditioned Krylov subspace iterative techniques, FMM accelerated matrix vector products, a novel FMM-based preconditioner, and fast translation techniques that enable us to achieve an overall algorithm in which the cost of the matrix-vector multiplication grows with N as $O(N \log N)$ and with the third power of the wave number. We discuss the convergence of the iterative techniques, selection of the truncation number, errors in the solution, and other issues. The results of the solution of test problems obtained with the method for $N \sim 10^2 - 10^4$ for different wave numbers are presented. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1853017]

PACS numbers: 43.20.Fn

Pages: 1744–1761

I. INTRODUCTION

Problems of acoustic and electromagnetic wave propagation through a medium consisting of the carrier fluid or solid and inclusions in the form of solid particles, droplets, and bubbles are of great practical and theoretical interest, and there are several approaches for the modeling of these. One approach is based on the development of continuum theories that treat a large system as a single complex medium with special properties that are obtained on the basis of spatial or other types of averaging (e.g., see Nigmatulin, 1990; Cafilisch *et al.*, 1985). This approach is validated against experiments and has many strengths, such as a relative simplicity of the description, a clear relation to the physics, and the operation with quantities that can be measured. However, as with any theory, the continuum theories have some limits and restrictions that follow from basic assumptions. For example, the models normally are based on assumptions that the medium is acoustically homogeneous (Gumerov *et al.*, 1988; Duraiswami and Prosperetti, 1995). This means that the wavelengths should be much larger than the characteristic size a and distance d between the particles ($ka \ll 1$, $kd \ll 1$, where k is the wave number). While some assumptions can be relaxed, an accurate averaging procedure should take into account many real effects, and unavoidably introduces assumptions that permit one to link the description of the processes at the “microlevel” with the behavior of quantities at the “macrolevel.”

Another approach is based on the direct computation of the acoustic field in the system at the level of particle sizes, based on fundamental governing equations applicable at the “microlevel.” While valid and free of many limitations of the continuum theory, this approach is computationally challenging. It is more or less obvious that there should be substantial efforts toward some theories that combine direct

computations, stochastic and continuum medium approach to extend their limits, and, in fact, this is not a purpose of this paper. Here we are interested with a solution of some “intermediate” class of problems, when, on the one hand, the number of particles is large enough to make conclusions of a statistical nature, and, on the other hand, may be small compared to real systems.

The problems, which are out of the range of continuum theories, are related to the acoustic wavelengths, which are comparable with the size of the particles. This may include both ultrasound propagation in dispersed systems, or lower-frequency sound propagation in some environment with larger objects (e.g., fish). In this case the number of scatterers in the system can be, say, several thousands. Since in the present paper we are solving the problem for the three-dimensional (3D) Helmholtz equation, which also appears in electromagnetics, the method may also be applicable to the solution of problems in diffraction optics, x-ray scattering, and so on. In many situations the scatterers can be assumed to be spherical, and the boundary conditions on the scatterer have a strong influence on the results.

There exist several approaches to the solution of the multiple scattering problems, and the T-matrix method and its modifications are among these approaches (Waterman and Truell, 1961; Peterson and Strom, 1974; Varadan and Varadan, 1980; Wang and Chew, 1993; Mischenko *et al.*, 1996; Koc and Chew, 1998) and recently in our paper [Gumerov and Duraiswami, 2002; which in the sequel we will refer to as (GD02)]. A solution of the problem for spherical scatterers based on a multipole re-expansion technique was presented in (GD02). This solution is based on a representation of the scattered field for each sphere by a series of multipoles, the coefficients of which can be determined by applying the impedance boundary conditions cor-

responding to each sphere. This was achieved using re-expansion matrices, which enable one to find the expansion coefficients in a basis conveniently located at the center of an arbitrary scatterer. Since the series are infinite, we used truncation to keep only p^2 terms in the series, where p is a quantity termed the truncation number.

In (GD02) the computational complexity for solution of this problem for N scatterers is $O(p^6 N^3)$. While we found that this method is reliable and fast for relatively small N and p , the solution of problems for a larger number of scatterers ($N \geq 10^3$) or higher frequencies (since for convergence $p > ka$) with this method is computationally intractable.

Such computations are also known as “brute-force” method, and there exist several papers in electromagnetics and acoustics that present algorithms of lower complexity. For example, Wang and Chew (1993), present a generalized recursive aggregate T-matrix algorithm (RATMA), with complexity $O(p^2 P^4 N)$ (in our notation), where P^2 is the number of harmonics to expand the field of the aggregated scatterers. The latter quantity is related to the size of the computational domain, D_0 , and $P \sim kD_0$. If we assume that the volume fraction of scatterers is constant, and so $D_0 \sim N^{1/3} a$, and since the truncation numbers can be estimated as $p \sim ka$, we can see that actually the RATMA scales as $O(p^6 N^{7/3})$. The RATMA does not improve the sixth power in the complexity dependence on p . That is why the computations with this algorithm can be performed only for low p (i.e., low ka). For example, Wang and Chew (1993) presented computations of electromagnetic scattering problem for large enough $N = 6859$, but for relatively low $kD_0 < 10$ (truncation numbers in these computations $p \leq 3$). In this context we should note that the accuracy of computations depends not only on the size of the domain or the size of scatterers (kD_0 and ka), but also on the interparticle distances, d/a . Even for $ka < 1$, truncation numbers can be, say $p \sim 15$, for the relatively low accuracy of computations of 1%, in the case that the scatterers are nearly touching (see the discussion on truncation numbers later in this paper).

Another, rather obvious, idea to improve the “brute-force” method is to use iterative methods for a solution of the linear system. If N_{iter} is the number of iterations to achieve a required accuracy and each iteration requires one matrix-vector multiplication, which can be performed directly for $O(p^4 N^2)$ operations, the total complexity of the algorithm will be $O(p^4 N^2 N_{\text{iter}})$. Comparing with the RATMA, we can see that for $N_{\text{iter}} \ll p^2 N^{1/3}$, an iterative technique is faster.

A further improvement of the algorithm complexity comes from the acceleration of the matrix-vector multiplication. Currently two major techniques are known to reduce the complexity of solution from $O(N^2)$ to $O(N \log N)$ or $O(N)$ (if we assume that the truncation number does not depend on N): the Fast Fourier Transform (FFT) and the Fast Multipole Method (FMM).

The FFT cannot be applied if the scatterers are not located in a regular grid. However, this can be fixed by using translations from the scatterer locations to the regular grid, which is an idea of the FMMFFT (Koc and Chew, 1998). As originally introduced, the FMMFFT utilizes the diagonal

forms of the translation operators and has complexity $O(p^4 N + p^2 N \log N)$. Recently some modifications of the FMMFFT, such as the FFTM (Ong *et al.*, 2004) appeared in the literature, which utilize different translation methods (recursively computed translation matrices) and have complexity $O(p^4 N \log N)$.

The FMM was introduced in Greengard and Rokhlin (1987) for solution of the Laplace equation in two and three dimensions. Later, this method was intensively studied and extended to a solution of many other problems. Sangani and Mo (1996) applied it to multiparticle systems using the analog of the T-matrix idea for the case of Laplace and Stokes interactions of multiple particles via the FMM.

In the original work, functions were represented via a set of expansion coefficients over a local or multipole basis. The translation operators in this case appeared to be dense matrices converting one set of the expansion coefficients to the other. If the number of expansion coefficients is p^2 , this results in $O(p^4)$ translation complexity and in $O(p^4 N)$ overall complexity of the algorithm if p is the same for all the levels of hierarchical space subdivision. In the case of the Laplace equation, this results in a fast enough algorithm. However, for the Helmholtz equation for a given accuracy, the number of terms needed in the expansion grows with the largest length scale in the problem. Consequently, if we want to use the method as is, then for the Helmholtz equation instead of p^2 we should use P^2 , which is the number of terms needed to represent the function at the coarsest level. For a uniform volume distribution of the scatterers with a constant volume fraction, we have $P \sim kD_0 \sim ka N^{1/3} \sim p N^{1/3}$, which results in $O(p^4 N^{7/3})$ complexity. This is slower than $O(p^4 N^2)$ required for direct matrix-vector multiplication, and so the FMM in its original form is not applicable to the scattering problem.

To reduce the cost of the FMM, an alternative function representation, via samples of the so-called “signature function” (also called the Herglotz wave function), was introduced by Rokhlin (1993). This representation gives rise to the diagonal forms of the translation operators. The number of samples is proportional to the number of harmonics in spectral representation, and here we can treat this quantity as p^2 . The diagonal forms provide $O(p^2)$ translation complexity. If the original FMM using the diagonal forms is employed for a solution of the Helmholtz equations, we should use P^2 instead of p^2 that results in $O(P^2 N) = O(p^2 N^{5/3})$ complexity of the matrix-vector multiplication. As was suggested by Rokhlin, the way to overcome these large complexities in the FMM is to vary the size of the vectors representing functions from level to level.

The algorithm of this type based on the diagonal forms of the translation operators with interpolation/interpolation procedures (MLFMA) is presented in Chew *et al.* (2001) and used for the acoustic scattering problems by Koc and Chew (1998). If the interpolation/interpolation procedures are performed with complexity $O(M)$ for $O(M)$ samples of the signature function with the required accuracy, then a solution of the Helmholtz equation can be obtained for $O(p^2 N)$ operations. The algorithm of Koc and Chew (1998), however, formally has complexity $O(p^4 N)$, since it requires a conver-

sion of the function representation from the set of multipole expansion coefficients to samples of the signature function, and further accounting for near interactions is also estimated (by us) as $O(p^3N)$. Our computations show that for dense systems a computation of the near interactions can be several times more expensive than a computation of the farfield interactions.

In fact, the $O(p^4N)$ conversion is not a necessary step for the algorithm if it operates with the expansion coefficients and uses matrix translation operators. The complexity of such nondiagonalized operations can be reduced to $O(p^3)$ for p^2 expansion coefficients using rotational-coaxial translation decomposition, and this is sufficient to achieve $O(p^3N \log N)$ overall complexity of the matrix-vector multiplication for the volume distribution of scatterers. While this fact is simple and we provide its proof in this paper, it seems to have been overlooked in the previous publications, where one can find statements that only the diagonal forms of the translation operators provide algorithms scaled as $O(N)$ or $O(N \log N)$ for the Helmholtz equation. The benefits of operating in the space of expansion coefficients also include stability of the translations, which does not require interpolation/interpolation or spherical filtering procedures that are necessary for the methods using diagonal forms and is an expensive part of it (Darve, 2000a; Chew *et al.*, 2001). Another benefit is the stability of computations for the cases of low ka and relatively large p .

While the paper of Koc and Chew (1998) presumably is the first that uses a combination of the T-matrix method and iterative techniques accelerated by the FMM and/or FFT for solution of acoustic scattering problems, several important issues still remain to be worked out. First, the iterative techniques presented in that paper are relatively standard, and their rate of convergence can be slow. As we found from our numerical experiments for systems of densely packed scatterers, N_{iter} at sufficiently large N may even grow as $O(N)$, which brings the total complexity of the FMM accelerated methods to $O(N^2)$ or $O(N^2 \log N)$. One of the ways to treat this problem is to use properly designed preconditioners. In this paper we introduce for the first time the flexible generalized minimal residual method (FGMRES) with a right dense preconditioner, which can also be computed with the FMM/GMRES. Our results show that this can substantially speed up the algorithm, and experimental data on the complexity fit complexities $O(N^{1+\beta})$ with low β . Second, we establish both *a priori* error bounds and used an *a posteriori* error check of the solution to see how the error depends on the relative scatterer locations and the truncation numbers used. Third, as it was presented, the T-matrix could be arbitrary, but at the same time the application of the method to real problems of interest in computational acoustics was limited. In the present paper we apply the solution to problems that have been studied extensively in acoustics, of multiple scattering from spheres. To solve this problem, we exploited the multipole solution for the boundary value problem for spheres presented in (GD02) and used the explicit T matrices, which depend on the scatterer sizes, impedances, etc. This allows one to model and investigate various parametric dependencies, and, move the method to the solution of gen-

eral practical acoustical problems. Finally, we used a variant of the FMM operating with expansion coefficients only, where the rectangularly truncated translation operators were computed using fast recursive algorithms and decompositions (Gumerov and Duraiswami, 2003). Numerical experiments show the performance and efficiency of our algorithm and support, that the matrix-vector multiplication is performed with $O(N)$ or $O(N \log N)$ complexity.

II. STATEMENT OF THE PROBLEM

The problem considered is that in (GD02), and is repeated here to establish the notation. Consider sound scattering by N spheres with radii a_1, \dots, a_N situated in \mathbb{R}^3 . The coordinates of the centers of the spheres are denoted as $\mathbf{r}'_q = (x'_q, y'_q, z'_q)$, $q = 1, \dots, N$. The scattering problem in the frequency domain is reduced to the solution of the Helmholtz equation for complex potential $\psi(\mathbf{r})$,

$$\nabla^2 \psi + k^2 \psi = 0, \quad (1)$$

with the following general impedance boundary conditions on the surface S_q of the q th sphere:

$$\left(\frac{\partial \psi}{\partial n} + i \sigma_q \psi \right) \Big|_{S_q} = 0, \quad q = 1, \dots, N, \quad (2)$$

where k is the wave number and σ_q are complex admittances, and $i = \sqrt{-1}$. In the particular case of sound-hard surfaces ($\sigma_q = 0$) we have the Neumann boundary conditions,

$$\partial \psi / \partial n \Big|_{S_q} = 0, \quad (3)$$

and in the case of sound soft surfaces ($\sigma_q = \infty$) we have the Dirichlet boundary conditions,

$$\psi \Big|_{S_q} = 0. \quad (4)$$

Usually the potential is represented in the form

$$\psi(\mathbf{r}) = \psi_{\text{in}}(\mathbf{r}) + \psi_{\text{scat}}(\mathbf{r}), \quad (5)$$

where $\psi_{\text{scat}}(\mathbf{r})$ is the potential of the scattered field. Far from the region occupied by spheres, the scattered field should satisfy the Sommerfeld radiation condition:

$$\lim_{r \rightarrow \infty} r \left(\frac{\partial \psi_{\text{scat}}}{\partial r} - ik \psi_{\text{scat}} \right) = 0. \quad (6)$$

We must determine $\psi(\mathbf{r})$ or $\psi_{\text{scat}}(\mathbf{r})$ at any \mathbf{r} on the surface of the spheres or outside them.

III. SOLUTION USING THE MULTIPOLE RE-EXPANSION METHOD

A. T matrix for a single scatterer

We first solve the scattering problem (1)–(6) for a single scatterer ($N = 1$) in an arbitrary incident field. We assume that the incident field has no singularities in the domain inside the scatterer or on its boundary, and since it satisfies the Helmholtz equation it can be expanded near the center of this scatterer, \mathbf{r}'_q ($q = 1$), in the form

$$\psi_{\text{in}}(\mathbf{r}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n E_n^{(q)m} R_n^m(\mathbf{r}_q) = \mathbf{E}^{(q)} \cdot \mathbf{R}(\mathbf{r}_q),$$

$$\mathbf{r}_q = \mathbf{r} - \mathbf{r}'_q. \quad (7)$$

Here $E_n^{(q)m}$ are the expansion coefficients of the potential over the basis of the regular solutions $R_n^m(\mathbf{r} - \mathbf{r}'_q)$ of the Helmholtz equation in the spherical coordinates $\mathbf{r}_q = (r_q, \theta_q, \varphi_q)$ connected with the center of the sphere:

$$R_n^m(\mathbf{r}_q) = j_n(kr_q) Y_n^m(\theta_q, \varphi_q),$$

$$\mathbf{r}_q = r_q (\sin \theta_q \cos \varphi_q, \sin \theta_q \sin \varphi_q, \cos \theta_q),$$

$$n = 0, 1, \dots, \quad m = -n, \dots, n, \quad (8)$$

where $j_n(kr)$ are the spherical Bessel functions of the first kind and $Y_n^m(\theta, \varphi)$ are the orthonormal spherical harmonics:

$$Y_n^m(\theta, \varphi) = (-1)^m \sqrt{\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} P_n^{|m|}(\cos \theta) e^{im\varphi},$$

$$n = 0, 1, \dots, \quad m = -n, \dots, n, \quad (9)$$

where $P_n^m(\mu)$ are the associated Legendre functions. Further, for compactness of notation, we use the following abbreviation to denote the sum represented by the product of two vectors: $\mathbf{A} = \{A_n^m\}$ and $\mathbf{B} = \{B_n^m\}$:

$$\mathbf{A} \cdot \mathbf{B} = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^m B_n^m, \quad \mathbf{A} = \{A_n^m\}, \quad \mathbf{B} = \{B_n^m\}. \quad (10)$$

The scattered field $\psi^{(q)}(\mathbf{r}) = \psi_{\text{scat}}(\mathbf{r})$ can be represented in the form

$$\psi^{(q)}(\mathbf{r}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^{(q)m} S_n^m(\mathbf{r}_q) = \mathbf{A}^{(q)} \cdot \mathbf{S}(\mathbf{r}_q). \quad (11)$$

Here $A_n^{(q)m}$ are the expansion coefficients of the potential over the basis of the singular at $\mathbf{r} = \mathbf{r}'_q$ radiating solutions $S_n^m(\mathbf{r} - \mathbf{r}'_q)$ of the Helmholtz equation:

$$S_n^m(\mathbf{r}_q) = h_n(kr_q) Y_n^m(\theta_q, \varphi_q),$$

$$n = 0, 1, \dots, \quad m = -n, \dots, n, \quad (12)$$

where $h_n(kr)$ are the spherical Hankel functions of the first kind.

By ‘‘T matrix for scatterer q ’’ we mean the matrix $\mathbf{T}^{(q)}$ with elements $T_{nn'}^{(q)mm'}$ that relates the expansion coefficients of the incident and scattered fields:

$$\mathbf{A}^{(q)} = \mathbf{T}^{(q)} \mathbf{E}^{(q)},$$

$$\left(A_n^{(q)m} = \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} T_{nn'}^{(q)mm'} E_{n'}^{(q)m'}, \right.$$

$$\left. n = 0, 1, \dots, \quad m = -n, \dots, n \right). \quad (13)$$

In general, the T matrix can be determined for a scatterer of an arbitrary shape by a solution of the corresponding boundary value problem (1)–(6). For the case of spheres, this solution can be obtained analytically, as shown in (GD02).

Based on this solution, the T matrix for a spherical scatterer of radius a_q and complex admittance σ_q is a diagonal matrix with elements

$$T_{nn'}^{(q)mm'} = - \frac{j_n'(ka_q) + (\sigma_q/k) j_n(ka_q)}{h_n'(ka_q) + (\sigma_q/k) h_n(ka_q)} \delta_{nn'} \delta_{mm'}, \quad (14)$$

where $n, n' = 0, 1, \dots$; $m = -n, \dots, n$; and $m' = -n', \dots, n'$.

B. Decomposition of the scattered field

Due to the linearity of the problem, the scattered field can be represented in the form

$$\psi_{\text{scat}}(\mathbf{r}) = \sum_{q=1}^N \psi^{(q)}(\mathbf{r}), \quad (15)$$

where $\psi^{(q)}(\mathbf{r})$ can loosely be thought of as the part of the scattered field introduced by the q th sphere, though, of course, it contains the influence of all the spheres. Each potential $\psi_q(\mathbf{r})$ is regular outside the q th sphere and satisfies the Sommerfeld radiation condition, so that it can be represented as prescribed by Eq. (11). In the vicinity of the q th scatterer we can write Eq. (15) as

$$\psi_{\text{scat}}(\mathbf{r}) = \psi^{(q)}(\mathbf{r}) + \phi^{(q)}(\mathbf{r}),$$

$$\phi^{(q)}(\mathbf{r}) = \sum_{q' \neq q} \psi^{(q')}(\mathbf{r}), \quad q, q' = 1, \dots, N. \quad (16)$$

The latter sum represents the function $\phi^{(q)}(\mathbf{r})$ that is regular in the part of the domain occupied by the q th scatterer and so can be represented as

$$\phi^{(q)}(\mathbf{r}) = \mathbf{B}^{(q)} \cdot \mathbf{R}(\mathbf{r}_q), \quad (17)$$

where $\mathbf{B}^{(q)} = \{B_n^{(q)m}\}$ is a vector of expansion coefficients over the basis of regular spherical functions (8). These coefficients can be expressed via the coefficients $\mathbf{A}^{(q')}$, $q' = 1, \dots, N$, $q' \neq q$ for all the scatterers, except of the q th scatterer, as

$$\mathbf{B}^{(q)} = \sum_{q' \neq q} (\mathbf{S}|\mathbf{R})(\mathbf{r}'_{q'}, \mathbf{r}'_q) \mathbf{A}^{(q')},$$

$$\mathbf{r}'_{q'} = \mathbf{r}'_q - \mathbf{r}'_{q'}, \quad q, q' = 1, \dots, N, \quad (18)$$

where $(\mathbf{S}|\mathbf{R})(\mathbf{r}'_{q'}, \mathbf{r}'_q)$ is the re-expansion matrix with elements $(S|R)_{nn'}^{mm'}(\mathbf{r}'_{q'}, \mathbf{r}'_q)$ defined as the coefficients of expansion of the singular basis functions centered at $\mathbf{r}'_{q'}$ over the regular basis functions centered at \mathbf{r}'_q (see Gumerov and Duraiswami, 2003),

$$S_{nn'}^{m'}(\mathbf{r}'_{q'}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n (S|R)_{nn'}^{mm'}(\mathbf{r}'_{q'}, \mathbf{r}'_q) R_n^m(\mathbf{r}_q),$$

$$q, q' = 1, \dots, N, \quad q' \neq q. \quad (19)$$

Now, using Eqs. (5), (7), (16), (11), and (17), we can represent the total field near the q th scatterer in the form

$$\psi(\mathbf{r}) = \mathbf{E}_{\text{eff}}^{(q)} \cdot \mathbf{R}(\mathbf{r}_q) + \mathbf{A}^{(q)} \cdot \mathbf{S}(\mathbf{r}_q), \quad \mathbf{E}_{\text{eff}}^{(q)} = \mathbf{E}^{(q)} + \mathbf{B}^{(q)}, \quad (20)$$

where we introduced the notation $\mathbf{E}_{\text{eff}}^{(q)}$ for the expansion coefficients of the “effective” incident field near the q th scatterer. Indeed, the first equation (20) shows that the multiple scattering problem is equivalent to a single scattering problem for the q th scatterer placed in the field, which is a superposition of the actual incident field and the fields scattered by the other scatterers. In this case the solution of the boundary value problem (1)–(6) for the q th scatterer is given by the \mathbf{T} matrix, which relates the expansion coefficients $\mathbf{A}^{(q)}$ and $\mathbf{E}_{\text{eff}}^{(q)}$ as prescribed by Eq. (13). So we have

$$\mathbf{A}^{(q)} = \mathbf{T}^{(q)}(\mathbf{E}^{(q)} + \mathbf{B}^{(q)}), \quad q = 1, \dots, N. \quad (21)$$

Now we can see that the system of equations (18) and (21) with respect to $\mathbf{A}^{(q)}$ and $\mathbf{B}^{(q)}$ is closed, as $\mathbf{T}^{(q)}$ and $\mathbf{E}^{(q)}$ are known (given). The coefficients $\mathbf{B}^{(q)}$ can be excluded to reduce the size of the system:

$$(\mathbf{T}^{(q)})^{-1} \mathbf{A}^{(q)} - \sum_{q' \neq q} (\mathbf{S}|\mathbf{R})(\mathbf{r}'_{q',q}) \mathbf{A}^{(q')} = \mathbf{E}^{(q)}, \quad q = 1, \dots, N. \quad (22)$$

This system also can be represented as a single equation,

$$\mathbf{L}\mathbf{A} = \mathbf{E}, \quad (23)$$

where the vectors and matrices are stacked as

$$\mathbf{L} = \begin{pmatrix} (\mathbf{T}^{(1)})^{-1} & -(\mathbf{S}|\mathbf{R})(\mathbf{r}'_{21}) & \cdots & -(\mathbf{S}|\mathbf{R})(\mathbf{r}'_{N1}) \\ -(\mathbf{S}|\mathbf{R})(\mathbf{r}'_{12}) & (\mathbf{T}^{(2)})^{-1} & \cdots & -(\mathbf{S}|\mathbf{R})(\mathbf{r}'_{N2}) \\ \cdots & \cdots & \cdots & \cdots \\ -(\mathbf{S}|\mathbf{R})(\mathbf{r}'_{1N}) & -(\mathbf{S}|\mathbf{R})(\mathbf{r}'_{2N}) & \cdots & (\mathbf{T}^{(N)})^{-1} \end{pmatrix}, \quad (24)$$

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}^{(1)} \\ \mathbf{A}^{(2)} \\ \cdots \\ \mathbf{A}^{(N)} \end{pmatrix}, \quad \mathbf{E} = \begin{pmatrix} \mathbf{E}^{(1)} \\ \mathbf{E}^{(2)} \\ \cdots \\ \mathbf{E}^{(N)} \end{pmatrix}.$$

This system can be solved by appropriately truncating the infinite vectors to p^2 coefficients ($\mathbf{A}^{(q)} = \{A_n^{(q)m}\}$, $n = 0, \dots, p-1$, $m = -n, \dots, n$). The final linear system with Np^2 unknowns is the same as in (GD02) [where Eq. (22) was multiplied by matrix $\mathbf{T}^{(q)}$ for each q]. Therefore $O(N^3 p^6)$ operations are required for a direct solution of this system. Using Eqs. (15) and (11) knowing all $\mathbf{A}^{(q)}$ enables the determination of $\psi_{\text{scat}}(\mathbf{r})$ everywhere in space. For the surface values of the potential and its normal derivatives simpler expressions can be used [see (GD02)].

IV. ITERATIVE METHODS

The first substantial reduction in complexity can be achieved by using iterative methods for solving the linear system. Assuming that each iteration involves one large dense matrix-vector multiplication of complexity $O(N^2 p^4)$ and that the iterative process converges to the required accuracy in N_{iter} steps, we immediately obtain a computational complexity $O(N_{\text{iter}} N^2 p^4)$ instead of $O(N^3 p^6)$. For example, if $N \sim 100$, $p \sim 10$, and $N_{\text{iter}} \sim 10$, the savings could be of order 10^3 times. Of course convergence may be slow, but as long as $N_{\text{iter}} \ll Np^2$, the iterative procedure brings substantial savings. We note in this context that there exist methods

(e.g., based on Krylov subspaces) that guarantee that an exact solution will be achieved in $N_{\text{iter}} \leq Np^2$ in exact arithmetics, while in practice the goal is to achieve this at much smaller N_{iter} .

Treating the multiple scattering problem as the solution of a dense linear system, we can employ different iterative methods starting from simple iterations to various versions of conjugate gradient, and Krylov subspace methods (see, e.g., Saad, 2003). Convergence of these methods can be improved by using well-designed preconditioners, block decompositions, etc.

The description and analysis of the entire area of iterative techniques is beyond the scope of this paper, and we refer the interested reader to the textbooks (Kelly, 1995; Saad, 2003). The use of a preconditioning matrix plays an important role, and the most efficient iterative schemes we used were based on the Flexible GMRES (FGMRES), which uses as the right preconditioner an approximate (GMRES) inverse of a matrix that approximates the system matrix \mathbf{L} (24). We describe details of this iterative technique in the next section since it is related closely to the FMM-based matrix-vector product.

Any iteration method builds a sequence of approximations to the solution

$$\mathbf{A}_0^{(q)}, \mathbf{A}_1^{(q)}, \dots; \quad \lim_{j \rightarrow \infty} \mathbf{A}_j^{(q)} = \mathbf{A}^{(q)}, \quad q = 1, \dots, N, \quad (25)$$

where $\mathbf{A}_0^{(q)}$ is some initial guess, e.g.,

$$\mathbf{A}_0^{(q)} = \mathbf{T}^{(q)} \mathbf{E}^{(q)}, \quad q = 1, \dots, N, \quad (26)$$

and stops when the convergence criterion is achieved. The convergence condition can be, e.g.,

$$\max_q \|\mathbf{A}_{j+1}^{(q)} - \mathbf{A}_j^{(q)}\| < \epsilon, \quad (27)$$

where $\|\cdot\|$ is some norm to measure the distance between two vectors (e.g., $\|\mathbf{A}\| = \max_{n,m} |A_n^m|$). Also, for each iteration it involves a computation of matrix vector product $\mathbf{L}\mathbf{A}_j$ for a given input vector \mathbf{A}_j . Thus, the speed of an iterative method can be improved by reducing the number of iterations and speeding up the embedded matrix vector products.

If we turn to the equation to be solved (23), we can see from Eq. (24) that the matrix \mathbf{L} can be decomposed as

$$\mathbf{L} = \mathbf{T}^{-1} - (\mathbf{S}|\mathbf{R}), \quad (28)$$

where \mathbf{T}^{-1} is a diagonal matrix for spherical scatterers and block-diagonal otherwise, and $(\mathbf{S}|\mathbf{R})$ is the matrix consisting of the $(\mathbf{S}|\mathbf{R})(\mathbf{r}'_{q,q'})$ blocks. Assuming that the product $\mathbf{T}^{-1}\mathbf{A}$ can be computed rapidly, we use the FMM to accelerate the computation of $\mathbf{B} = (\mathbf{S}|\mathbf{R})\mathbf{A}$, or

$$\mathbf{B}_j^{(q)} = \sum_{q' \neq q} (\mathbf{S}|\mathbf{R})(\mathbf{r}'_{q',q}) \mathbf{A}_j^{(q')}, \quad q = 1, \dots, N, \quad j = 0, 1, \dots \quad (29)$$

V. FAST MULTIPOLE METHOD

The second substantial speed up of the solution can be obtained by accelerating the matrix-vector product (23) in the iterative methods via the multilevel Fast Multipole

Method (FMM). We refer the reader to details of the method in Greengard (1988) and for particulars of the method we used (Gumerov *et al.*, 2003). Below we describe the method used for the multiple scattering problems.

A. Data structures

The FMM operates with the spatial data organized in octrees. This means that the computational domain is the cube, or box, of size $D_0 \times D_0 \times D_0$ that encloses all the scatterers. We assign level 0 to this cube. Further, the computational domain is subdivided into eight smaller boxes $D_1 \times D_1 \times D_1$, each of size $D_1 = 0.5D_0$. This is referred to as level 1. Level 2 is obtained by the subdivision of boxes in level 1 by 8, and so on until some level l_{\max} , where there are $8^{l_{\max}}$ boxes of size $D_{l_{\max}} = 2^{-l_{\max}}D_0$. To each box we assign some index (n, l) , $n = 0, \dots, 8^l - 1$, $l = 0, \dots, l_{\max}$. Such a space partitioning generates a hierarchical data structure with “parent–children” relationships, where any box of level $l = 0, \dots, l_{\max} - 1$ can be considered as a “parent,” with 8 children boxes obtained by its subdivision. Neighbors of a box B are boxes at the same level, that share at least one boundary point with box B . Boxes that are not adjacent to the boundaries of the computational domain have 26 neighbors, which together with the box itself form the “neighborhood” of the box.

In the multiple scattering problem we say that scatterer q belongs to box (n, l) if the center \mathbf{r}'_q of the smallest sphere enclosing the scatterer is located inside that box (for spherical scatterers this sphere is simply its surface). Any box may intersect a scatterer, which does not belong to it (see Fig. 1). However, we require that the size of the smallest box (at the maximum level of space partitioning) is determined as below to ensure both the correctness of the FMM procedure and ensure the validity of the multipole re-expansions on the surfaces of neighboring scatterers. If the scatterers q and q' belong to two different boxes at the finest level with indices (n_q, l_{\max}) and $(n_{q'}, l_{\max})$, which are not neighbors, then the field scattered by q' can be described by the local expansion of type (17) centered at the center of box (n_q, l_{\max}) in the domain occupied by the sphere surrounding the scatterer q . The maximum distance from the center of box $(n_{q'}, l_{\max})$ to a point on the surface of scatterer q' is $D_{l_{\max}}\sqrt{3}/2 + a_{\max}$, where a_{\max} is the maximum possible size of the scatterer, $a_{\max} = \max(a_1, \dots, a_N)$. The minimum possible distance from the center of the box $(n_{q'}, l_{\max})$ to a point on the surface of the q th scatterer is then $3D_{l_{\max}}/2 - a_{\max}$. The requirement on the validity of expansions will be satisfied if $3D_{l_{\max}}/2 - a_{\max} > D_{l_{\max}}\sqrt{3}/2 + a_{\max}$, which yields the following condition:

$$D_{l_{\max}} > \frac{4}{3 - \sqrt{3}} a_{\max} \approx 3.15 a_{\max},$$

$$l_{\max} < \log_2 \left(\frac{3 - \sqrt{3}}{4} \frac{D_0}{a_{\max}} \right). \quad (30)$$

Note that this condition also ensures that the farfield re-expansion about the center of the box (n_q, l_{\max}) of the field

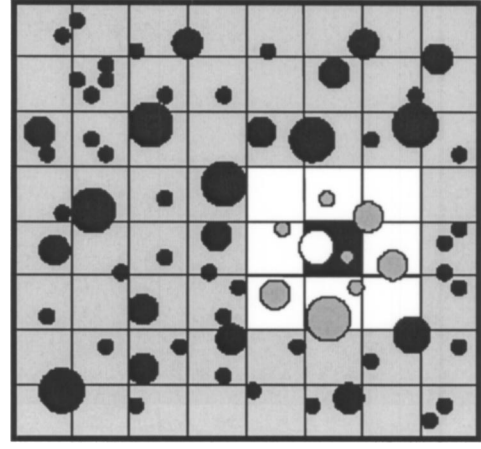


FIG. 1. Illustration of space partitioning and data structure for multiple scattering problem. The minimal spheres enclosing each scatterer are shown (so an actual scatterer of arbitrary shape is located inside the sphere). The white sphere belongs to the black box. The gray spheres are its neighbors. These are spheres which belong (in terms of their centers) to the E_2 neighborhood of the black box. The black spheres are the far spheres, they belong to the E_3 neighborhood of the black box (the centers are located outside the E_2 neighborhood). For a description of different neighborhoods see also Fig. 2.

scattered by scatterer q is valid at any point on the surface of the scatterer q' .

A stronger limitation on l_{\max} follows from formal requirements of the translation theory. Again, we can consider two closest boxes $(n_{q'}, l_{\max})$ and (n_q, l_{\max}) , which are not neighbors, and try to evaluate the effect of scatterer q' on q . The multipole series representing the potential of the scatterer q' , whose center in the worst case is located at distance $D_{l_{\max}}\sqrt{3}/2$ from the center of the box $(n_{q'}, l_{\max})$ should be translated first to the center of this box, then to the center of the box (n_q, l_{\max}) , and finally to the center of scatterer q with a guarantee that the final local expansion is valid within the sphere of radius a_{\max} . This can be provided if $2D_{l_{\max}} > D_{l_{\max}}\sqrt{3}/2 + D_{l_{\max}}\sqrt{3}/2 + a_{\max}$, which is a condition that spheres of radii $D_{l_{\max}}\sqrt{3}/2$ and $D_{l_{\max}}\sqrt{3}/2 + a_{\max}$ circumscribing the boxes $(n_{q'}, l_{\max})$ and (n_q, l_{\max}) , respectively, do not intersect ($2D_{l_{\max}}$ is the distance between the box centers). This condition can also be rewritten as

$$D_{l_{\max}} > \frac{1}{2 - \sqrt{3}} a_{\max} \approx 3.73 a_{\max},$$

$$l_{\max} < \log_2 \left[(2 - \sqrt{3}) \frac{D_0}{a_{\max}} \right]. \quad (31)$$

A further limitation on l_{\max} should be imposed, since we use truncated expansions. The truncation error is a function of the truncation number, which in turn depends on the distance between the center of the domain of local expansion and the closest point on the sphere surrounding the box containing the multipole, $b_{l_{\max}} = 2D_{l_{\max}} - D_{l_{\max}}\sqrt{3}/2$. The dimensionless parameter,

$$\delta_{\min} = b_{l_{\max}} / (D_{l_{\max}}\sqrt{3}/2 + a_{\max}) > 1, \quad (32)$$

then plays an important role for a selection of the truncation number (see the section below) ($p \rightarrow \infty$ for $\delta_{\min} \rightarrow 1$). If this parameter is specified, then we have the following modification for condition (31):

$$D_{l_{\max}} \geq \frac{2\delta_{\min}}{4 - (1 + \delta_{\min})\sqrt{3}} a_{\max},$$

$$l_{\max} \leq \log_2 \left[\frac{4 - (1 + \delta_{\min})\sqrt{3}}{2\delta_{\min}} \frac{D_0}{a_{\max}} \right], \quad \delta_{\min} > 1. \quad (33)$$

It can be seen that in any case $\delta_{\min} < 4/\sqrt{3} - 1 \approx 1.31$. Note that in some test cases we found numerically that the algorithm may continue to perform satisfactorily when l_{\max} exceeds the value prescribed by conditions (30)–(33).

B. Field decomposition

The field $\phi^{(q)}(\mathbf{r})$ [see Eq. (17)] due to the potentials centered at spheres other than the q th sphere can be decomposed into two parts:

$$\phi^{(q)}(\mathbf{r}) = \phi_{\text{near}}^{(q)}(\mathbf{r}) + \phi_{\text{far}}^{(q)}(\mathbf{r}), \quad \mathbf{B}^{(q)} = \mathbf{B}_{\text{near}}^{(q)} + \mathbf{B}_{\text{far}}^{(q)},$$

$$\phi_{\text{near}}^{(q)}(\mathbf{r}) = \mathbf{B}_{\text{near}}^{(q)} \cdot \mathbf{R}(\mathbf{r}_q), \quad \phi_{\text{far}}^{(q)}(\mathbf{r}) = \mathbf{B}_{\text{far}}^{(q)} \cdot \mathbf{R}(\mathbf{r}_q), \quad (34)$$

where $\phi_{\text{near}}^{(q)}(\mathbf{r})$ is the field potential centered at the spheres that belong to the neighborhood $\Omega^{(q)}$ of the box containing the q th sphere at level l_{\max} , $\phi_{\text{far}}^{(q)}(\mathbf{r})$ is the field scattered by the other spheres, and $\mathbf{B}_{\text{near}}^{(q)}$ and $\mathbf{B}_{\text{far}}^{(q)}$ are the respective expansion coefficients in the basis $\mathbf{R}(\mathbf{r}_q)$. According to Eq. (18), these coefficients can be written in the form

$$\mathbf{B}_{\text{near}}^{(q)} = \sum_{\mathbf{r}_{q'} \in \Omega^{(q)}, q' \neq q} (\mathbf{S}|\mathbf{R})(\mathbf{r}'_{q'}) \mathbf{A}^{(q')},$$

$$\mathbf{B}_{\text{far}}^{(q)} = \sum_{\mathbf{r}_{q'} \notin \Omega^{(q)}} (\mathbf{S}|\mathbf{R})(\mathbf{r}'_{q'}) \mathbf{A}^{(q')}. \quad (35)$$

In the iterative algorithm, the coefficients $\mathbf{A}^{(q)}$ are known from the previous iteration step, and the purpose is to determine $\mathbf{B}^{(q)}$ for the next iteration step [see Eq. (29)]. As there are not many scatterers in the neighborhood of the q th sphere, the coefficients $\mathbf{B}_{\text{near}}^{(q)}$ can be computed directly using the first equation in (35). The major computational expense comes from the second sum. To compute $\mathbf{B}_{\text{far}}^{(q)}$ we use the FMM, as described in the next section. Based on a given set of coefficients $\{\mathbf{A}^{(q)}\}$, the FMM provides coefficients $\mathbf{D}^{(q)}$ of the farfield expansion near the center of the box at the finest level, containing the q th scatterer, $\mathbf{r}_*^{(q)}$:

$$\phi_{\text{far}}^{(q)}(\mathbf{r}) = \mathbf{D}^{(q)} \cdot \mathbf{R}(\mathbf{r} - \mathbf{r}_*^{(q)}), \quad q = 1, \dots, N. \quad (36)$$

To get $\mathbf{B}_{\text{far}}^{(q)}$, which are the coefficients for expansion about the center of the q th sphere we need to translate this expansion to the proper center. This can be done by using the local-to-local translation operator, or by applying the $(\mathbf{R}|\mathbf{R})$ re-expansion matrix to coefficients $\mathbf{D}^{(q)}$:

$$\mathbf{B}_{\text{far}}^{(q)} = (\mathbf{R}|\mathbf{R})(\mathbf{r}'_q - \mathbf{r}_*^{(q)}) \mathbf{D}^{(q)}, \quad q = 1, \dots, N. \quad (37)$$

Here the $(\mathbf{R}|\mathbf{R})$ re-expansion matrix has entries $(R|R)_{nn'}^{mm'}$ that are the expansion coefficients of the regular basis func-

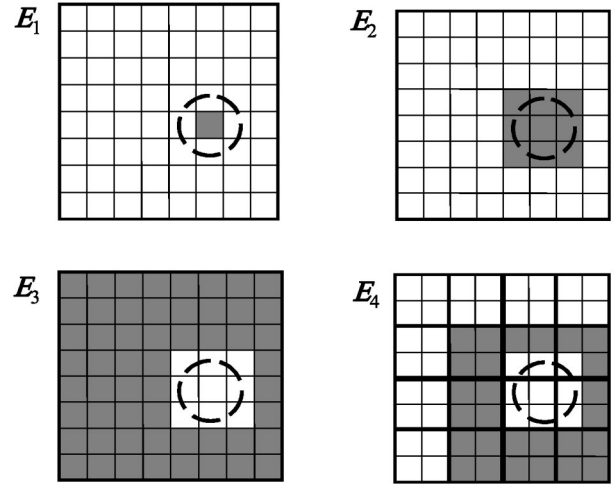


FIG. 2. The domains used for the construction of a hierarchical re-expansion procedure in FMM (shown in projection to a plane). A circle separates the box for which domains are drawn.

tion over the same type of basis, whose center is shifted:

$$R_{n'}^{m'}(\mathbf{r} - \mathbf{r}_*^{(q)}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n (R|R)_{nn'}^{mm'}(\mathbf{r}'_q - \mathbf{r}_*^{(q)}) R_n^m(\mathbf{r}_q),$$

$$q = 1, \dots, N, \quad n' = 0, 1, \dots, \quad m' = -n', \dots, n'. \quad (38)$$

C. Details of the fast multipole method

In the FMM we introduce the following hierarchical domains, associated with each level of space subdivision: $E_1(n, l)$, the set of points located inside box (n, l) ; $E_2(n, l)$, the set of points located inside the neighborhood of box (n, l) ; $E_3(n, l)$, the set of points located in the computational domain, which are outside the neighborhood of box (n, l) [this domain is complementary to $E_2(n, l)$], and $E_4(n, l)$, the set of points located in domain $E_2(\text{Parent}(n), l-1)$, but outside domain $E_2(n, l)$. For boxes, whose parents are not adjacent to the boundaries of the computational domain, $E_4(n, l)$, consists of 189 boxes of level l . These domains are shown in Fig. 2. We associate fields $\psi_j^{(n, l)}(\mathbf{r})$ with every domain $E_j(n, l)$, $j = 1, \dots, 4$. These are fields scattered by all the spheres, which belong to the respective domains. We also associate truncation numbers $p_2, \dots, p_{l_{\max}}$ with each level of subdivision $l = 2, \dots, l_{\max}$. The selection of these truncation numbers is based on the error analysis and is discussed later. The algorithm consists of two major parts. In the first part we set up the hierarchical data structure, based on the location and sizes of the scatterers. This step need be performed only once for a given scatterer configuration, and is not repeated during the iterative process. The second part consists of the upward pass and the downward pass, followed by the final summation step, which is described by Eqs. (35) and (37), leads to a determination of the expansion coefficients $\mathbf{B}^{(q)}$, $q = 1, \dots, N$.

1. Upward pass

The purpose of the upward pass is to determine coefficients $\mathbf{C}^{(n,l)}$ of expansion for all functions $\psi_1^{(n,l)}(\mathbf{r})$ over the bases of the S functions centered at the box centers $\mathbf{r}_*^{(n,l)}$:

$$\begin{aligned} \psi_1^{(n,l)}(\mathbf{r}) &= \mathbf{C}^{(n,l)} \cdot \mathbf{S}(\mathbf{r} - \mathbf{r}_*^{(n,l)}), \quad l=2, \dots, l_{\max}, \\ \forall n, \quad \mathbf{r}'_q &\in E_1(n,l), \quad q=1, \dots, N. \end{aligned} \quad (39)$$

Here the length of the truncated coefficient vectors $\mathbf{C}^{(n,l)}$ and, respectively, the basis functions $\mathbf{S}(\mathbf{r} - \mathbf{r}_*^{(n,l)})$ depends on the level l and is of length p_l^2 . To build the coefficients we apply the multipole-to-multipole, or the $(\mathbf{S}|\mathbf{S})$ translation operator, which can be thought of as a rectangularly truncated matrix with entries $(S|S)_{nn'}^{mm'}$, defined as re-expansion coefficients for the S functions:

$$\begin{aligned} S_{n'}^{m'}(\mathbf{r} + \mathbf{t}) &= \sum_{n=0}^{\infty} \sum_{m=-n}^n (S|S)_{nn'}^{mm'}(\mathbf{t}) S_n^m(\mathbf{r}), \\ n' &= 0, 1, \dots, \quad m' = -n', \dots, n', \end{aligned} \quad (40)$$

where \mathbf{t} is the translation vector determined by the location of the expansion centers. Note that for the Helmholtz equation we have $(\mathbf{S}|\mathbf{S})(\mathbf{t}) = (\mathbf{R}|\mathbf{R})(\mathbf{t})$.

- Step 1. For each scatterer determine the coefficients of the S expansion centered at the center of the box $\mathbf{r}_*^{(q)} = \mathbf{r}_*^{(n, l_{\max})}$ at the finest level, (n, l_{\max}) , to which this scatterer belongs using the multipole-to-multipole $S|S$ translation of the coefficients $\mathbf{A}^{(q)}$ to the center of this box:

$$\begin{aligned} \tilde{\mathbf{C}}^{(q)} &= (\mathbf{S}|\mathbf{S})(\mathbf{r}_*^{(q)} - \mathbf{r}'_q) \mathbf{A}^{(q)}, \quad q=1, \dots, N, \\ \mathbf{r}'_q, \mathbf{r}_*^{(q)} &\in E_1(n, l_{\max}). \end{aligned} \quad (41)$$

This step of the algorithm provides a representation of the scattered field for the q th scatterer (11) about the center of the box to which it belongs:

$$\psi^{(q)}(\mathbf{r}) = \mathbf{A}^{(q)} \cdot \mathbf{S}(\mathbf{r}_q) = \tilde{\mathbf{C}}^{(q)} \cdot \mathbf{S}(\mathbf{r} - \mathbf{r}_*^{(q)}). \quad (42)$$

Note that the length of truncated vectors $\mathbf{A}^{(q)}$ is p^2 , while the length of the truncated vectors $\tilde{\mathbf{C}}^{(q)}$ is $p_{l_{\max}}^2$. These lengths can be the same or different. We discuss this issue later.

Step 2. Consolidate all the expansions at the finest level, to obtain the expansion coefficients $\mathbf{C}^{(n, l_{\max})}$ of functions $\psi_1^{(n, l_{\max})}(\mathbf{r})$ over the S basis:

$$\mathbf{C}^{(n, l_{\max})} = \sum_{\mathbf{r}_*^{(q)} \in E_1(n, l_{\max})} \tilde{\mathbf{C}}^{(q)}, \quad \mathbf{r}_*^{(n, l_{\max})} = \mathbf{r}_*^{(q)}. \quad (43)$$

Step 3. For $l = l_{\max} - 1, \dots, 2$, recursively obtain expansion coefficients $\mathbf{C}^{(n,l)}$ for all other functions $\psi_1^{(n,l)}(\mathbf{r})$ over the S basis using the multipole-to-multipole translation operators from the centers of children boxes containing the scatterers, $\mathbf{r}_*^{(n', l+1)}$, to the center of their parent box, $\mathbf{r}_*^{(n,l)}$, and the consolidation of the coefficients:

$$\mathbf{C}^{(n,l)} = \sum_{n' \in \text{Children}(n,l)} (\mathbf{S}|\mathbf{S})(\mathbf{r}_*^{(n,l)} - \mathbf{r}_*^{(n', l+1)}) \mathbf{C}^{(n', l+1)}. \quad (44)$$

Here $\text{Children}(n,l)$ denotes the set of children boxes for box (n,l) . These boxes are of a smaller size than (n,l) and located at the level $l+1$.

2. Downward pass

The purpose of the downward pass is to determine coefficients $\mathbf{D}^{(n,l)}$ of expansion for all functions $\psi_3^{(n,l)}(\mathbf{r})$ over the bases of the R functions centered at $\mathbf{r}_*^{(n,l)}$:

$$\begin{aligned} \psi_3^{(n,l)}(\mathbf{r}) &= \mathbf{D}^{(n,l)} \cdot \mathbf{R}(\mathbf{r} - \mathbf{r}_*^{(n,l)}), \quad l=2, \dots, l_{\max}, \\ \forall n, \quad \mathbf{r}'_q &\in E_1(n,l), \quad q=1, \dots, N. \end{aligned} \quad (45)$$

Here the length of truncated coefficient vectors $\mathbf{D}^{(n,l)}$ and, respectively, the basis $\mathbf{R}(\mathbf{r} - \mathbf{r}_*^{(n,l)})$ depends on the level l and is p_l^2 . To build them we apply the multipole-to-local and local-to-local, or the $(\mathbf{S}|\mathbf{R})$ and $(\mathbf{R}|\mathbf{R})$ translation operators, which can be thought of as square and rectangularly truncated matrices, respectively, since the $(\mathbf{S}|\mathbf{R})$ translation is applied for boxes at the same level, while the $(\mathbf{R}|\mathbf{R})$ translation is used to translate coefficients from a lower level to a higher level. In the process we also determine coefficients $\tilde{\mathbf{D}}^{(n,l)}$ of expansion for potential $\psi_4^{(n,l)}(\mathbf{r})$ over the local bases:

$$\begin{aligned} \psi_4^{(n,l)}(\mathbf{r}) &= \tilde{\mathbf{D}}^{(n,l)} \cdot \mathbf{R}(\mathbf{r} - \mathbf{r}_*^{(n,l)}), \quad l=2, \dots, l_{\max}, \\ \forall n, \quad \mathbf{r}'_q &\in E_1(n,l), \quad q=1, \dots, N. \end{aligned} \quad (46)$$

We need to do this in terms of organization of the hierarchical procedure, which is based on the following property of the domains $E_3(n,l)$ and $E_4(n,l)$ (see Fig. 2):

$$E_3(n,l) = E_3(n', l-1) \cup E_4(n,l), \quad n' = \text{Parent}(n). \quad (47)$$

This results in

$$\psi_3^{(n,l)}(\mathbf{r}) = \psi_3^{(n', l-1)}(\mathbf{r}) + \psi_4^{(n,l)}(\mathbf{r}), \quad n' = \text{Parent}(n), \quad (48)$$

and enables recursion with respect to the levels of the octree.

- Step 1. (This step and step 2 of the downward pass are performed recursively for $l=2, \dots, l_{\max}$.) Obtain coefficients $\tilde{\mathbf{D}}^{(n,l)}$ for all boxes containing the scatterers. We obtain these coefficients by the multipole-to-local translations of respective coefficients $\mathbf{C}^{(n', l)}$, followed by the consolidation of the expansions:

$$\begin{aligned} \tilde{\mathbf{D}}^{(n,l)} &= \sum_{\mathbf{r}_*^{(n', l)} \in E_4(n,l)} (\mathbf{S}|\mathbf{R})(\mathbf{r}_*^{(n,l)} - \mathbf{r}_*^{(n', l)}) \mathbf{C}^{(n', l)}, \\ \forall n, n', \quad \mathbf{r}'_q &\in E_1(n,l), \quad \mathbf{r}'_{q'} \in E_1(n', l), \quad q, q' = 1, \dots, N. \end{aligned} \quad (49)$$

Step 2. At level $l=2$ we have $E_3(n,2) = E_4(n,2)$, which results in $\mathbf{D}^{(n,2)} = \tilde{\mathbf{D}}^{(n,2)}$. For levels $l=3, \dots, l_{\max}$ we have, from Eq. (48),

$$\mathbf{D}^{(n,l)} = (\mathbf{R}|\mathbf{R})(\mathbf{r}_*^{(n,l)} - \mathbf{r}_*^{(n', l-1)}) \mathbf{D}^{(n', l-1)} + \tilde{\mathbf{D}}^{(n,l)}, \quad (50)$$

$$n' = \text{Parent}(n), \quad l = 2, \dots, l_{\max}, \quad \forall n, \quad \mathbf{r}'_q \in E_1(n, l),$$

$$q = 1, \dots, N.$$

3. Final summation

The final summation step results in a determination of the scattered field at the current iteration, since the domains $E_3(n, l)$ and $E_2(n, l)$ are complementary. This step of the usual FMM procedure is skipped here, since all we need are the expansion coefficients at the finest level $\mathbf{D}^{(q)} = \mathbf{D}^{(n, l_{\max})}$, $\mathbf{r}'_q \in E_1(n, l_{\max})$, $q = 1, \dots, N$. Equation (37) then provides coefficients $\mathbf{B}_{\text{far}}^{(q)}$.

D. Truncation numbers

To apply the algorithm we need to specify the truncation numbers p for the coefficients $\{\mathbf{A}^{(q)}\}$ and truncation numbers $p_2, \dots, p_{l_{\max}}$. Their selection depends on the acceptable error of computations, which is determined by the rate of convergence of the series being truncated. Consider first the selection of p .

The series under consideration converge absolutely and uniformly only for $p \geq ka$. For relatively low ka ($ka \leq 10$), one can select p as

$$p = [ka] + p_0(ka, \epsilon, \delta), \quad p_0 \geq 1, \quad \delta = b/a \geq 1, \quad \epsilon > 0, \quad (51)$$

where p_0 depends on the prescribed accuracy ϵ and the distance b between the center of the scatterer and the closest point on the surface of the closest scatterer. While there are some theories for the expansion errors (e.g., Koc *et al.*, 1999; Darve, 2000b), they usually deal with the error of expansion of a single source. We suggest the evaluation of p_0 based on error maps (or their approximations) obtained as a solution of the scattering problem for two spheres placed in the field of the plane incident wave, $\psi_{\text{in}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}$. This problem can be solved very rapidly based on direct matrix inversion for coaxial spheres (GD02), where the parameters p_0 , ka , and δ can be varied within the range under consideration, while the error of the solution ϵ can be determined *a posteriori*, as the error in boundary conditions.

Indeed, we note that since any radiating basis function satisfies the Helmholtz equation, the expansion truncated with an arbitrary p will satisfy the Helmholtz equation and radiation condition (6). So the only equation that is not satisfied exactly by the approximate solution is the boundary condition (2). Once an approximate solution is computed, i.e., the expansion coefficients $\{\mathbf{A}^{(q)}\}$ are found, we can compute the normal derivative on the surface for any scatterer using the differentiation theorem for the spherical basis functions in an arbitrary direction specified by the unit vector $\mathbf{n} = (n_x, n_y, n_z)$ (Gumerov and Duraiswami, 2003):

$$\begin{aligned} \frac{1}{k} \mathbf{n} \cdot \nabla S_n^m(\mathbf{r}) &= \frac{1}{2} (n_x - i n_y) [b_{n+1}^{-m-1} S_{n+1}^{m+1}(\mathbf{r}) - b_n^m S_{n-1}^{m+1}(\mathbf{r})] \\ &+ \frac{1}{2} (n_x + i n_y) [b_{n+1}^{m-1} S_{n+1}^{m-1}(\mathbf{r}) - b_n^{-m} S_{n-1}^{m-1}(\mathbf{r})] \\ &+ n_z [a_{n-1}^m S_{n-1}^m(\mathbf{r}) - a_n^m S_{n+1}^m(\mathbf{r})], \end{aligned} \quad (52)$$

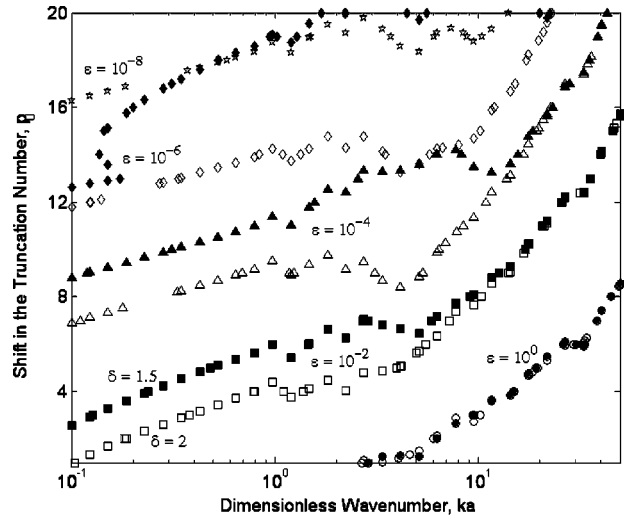


FIG. 3. Dependences of the shift in the truncation number, p_0 , defined by Eq. (51) on the dimensionless wave number ka for two sound-hard spheres of equal size. The curves are computed to provide the error $\epsilon^{(bc)}$ in boundary conditions shown near the curves. Two sets of curves marked by the filled and empty symbols correspond to different distances between the spheres ($\delta=1.5$ and $\delta=2$, respectively).

where the coefficients a_n and b_n are specified by Eqs. (62) and (63). Therefore, as the truncated solution is computed, the left-hand side of Eq. (2) can be determined, and there are no additional errors for derivative computation or approximation. If we sample the surface with M points, we can then compute the following errors in boundary conditions:

$$\epsilon_m^{(bc)} = \frac{M |\Delta_{bc}(\mathbf{y}_m)|}{[\sum_m |\psi|_{S_q}(\mathbf{y}_m)|^2]^{1/2}}, \quad \epsilon^{(bc)} = \max_m \epsilon_m^{(bc)}, \quad (53)$$

$$\Delta_{bc}(\mathbf{y}_m) = a_q \left(\frac{\partial \psi}{\partial n}(\mathbf{y}_m) + i \sigma_q \psi(\mathbf{y}_m) \right) \Big|_{S_q}, \quad m = 1, \dots, M.$$

Figure 3 illustrates dependencies $p_0(ka, \epsilon^{(bc)}, \delta)$ for two fixed values of δ and a range of ka . The spheres are sound-hard, $\sigma_q = 0$, $q = 1, 2$. Here for each curve $\epsilon^{(bc)}$ is fixed (p_0 is not an integer due to interpolation). $\epsilon^{(bc)}$ was computed according to Eq. (53), where each sphere was sampled with 181 equispaced points with respect to the spherical polar angle θ . For the low-frequency region, $ka \leq 1$ and $\epsilon^{(bc)} \geq 10^{-4}$, we have $p_0 \sim p_{00}(\epsilon, \delta) + \log(ka)$, which in the semi-logarithmic coordinates is a linear dependence. For higher frequencies, $ka \geq 1$, we have $p_0 \sim (ka)^\nu$, where theoretically $\nu = 1/3$ and that is pretty close to our computations. The dependence on δ in this range is rather weak (when δ is not very close to unity), while for lower frequencies the truncation number substantially depends on this parameter. Figure 4 illustrates dependencies $p_0(ka, \epsilon^{(bc)}, \delta)$ for fixed $ka = 1$ and varying δ . It is seen that if the spheres are very close to each other ($\delta \sim 1$), the truncation number should be high to provide small errors. However, if it is sufficient to compute a solution with, say, $\sim 1\%$ of relative error, these computations can be achieved with $p_0 \sim 10$, even when the spheres touch each other.

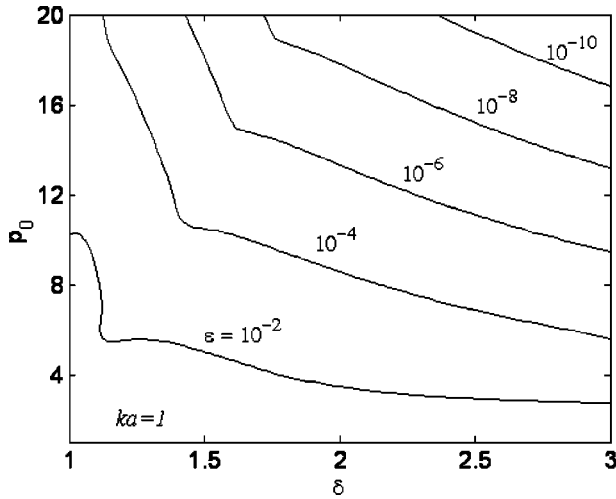


FIG. 4. Dependences of the shift in the truncation number, p_0 , defined by Eq. (51) on the dimensionless parameter δ characterizing the distance between the spheres for two sound-hard spheres of equal size. The curves are computed to provide the error $\epsilon^{(bc)}$ in boundary conditions shown near the curves.

For larger ka the truncation number can be selected based on the high-frequency asymptotics of the spherical Bessel functions (e.g., see Chew *et al.*, 2001):

$$p = P(ka, \epsilon) = ka + \frac{1}{2} \left(3 \ln \frac{1}{\epsilon} \right)^{2/3} (ka)^{1/3}. \quad (54)$$

In this equation the truncation number does not depend on δ , while such a dependence can be introduced. Equation (54) is qualitatively consistent with the behavior of the error observed in computations where δ is not very close to 1. To relate it to the actual error observed for a solution of the two sphere scattering problem one should relate ϵ to $\epsilon^{(bc)}$, and we found that for the range plotted in Fig. 3 one can set $\epsilon \sim 10^{-2} \epsilon^{(bc)}$ to fit results for $ka > 1$.

Similar formulas can be applied for a computation of the truncation numbers used in the FMM, in which case we need to specify what we mean by a and δ here. The parameter a can be selected as the radius of the smallest sphere that encloses the box of level l plus a_{\max} and δ formally is limited by the value $\delta < 1.31$ (see the discussion on l_{\max} previously, while a more accurate analysis is required here, since the “worst case” analysis usually substantially overestimates actual errors). Since the size of this box is $D_l = 2^{-l} D_0$, one can set

$$a_l = 2^{-l-1} 3^{1/2} D_0 + a_{\max}, \quad p_l = P(ka_l, \epsilon). \quad (55)$$

In computations we used this formula for automatically setting the level-dependent truncation numbers with function $P(ka, \epsilon)$ specified by Eq. (54), where ϵ was some prescribed error. In computations we also performed an *a posteriori* actual error check using Eq. (53), which is valid for an arbitrary number of spheres.

We note also that for the present problem, when the number of scatterers can be large, and so for fixed ka and the volume fraction of scatterers we have $D_0 \sim N^{1/3} a$, truncation numbers p_l used in the FMM can be substantially larger than p . In other words, if the wavelength is comparable with the

size of the scatterer, then the size of the computational domain is much larger than the wavelength for large N .

We can note that in the case when the spheres are close or may touch each other, p can be larger than p_l to provide the required accuracy. To avoid unnecessary computations in this case, translations for the farfield expansions at the finest level can be made to obtain only $p_{l_{\max}}^2$ terms, since terms of degree n corresponding to $p_{l_{\max}} \leq n < p$ are due to the influence of close neighbors. Moreover, in computations of the near field, we can subdivide the set of neighbor spheres into two sets: those that are closer than some prescribed b_* (or δ_*), for which the interaction should be computed using all p^2 terms, and the other neighbors, whose interactions with the given sphere can be efficiently computed using only $p_{l_{\max}}^2$ terms. We found numerically that these tricks may bring substantial savings in computational time, while almost not affecting the accuracy of the procedure.

E. Complexity estimates

In the FMM, translations occur between two subsequent levels l and $l+1$ (for the local-to-local or multipole-to-multipole translations) or on the same level l (for the multipole-to-local translations). Since the number of coefficients representing the potential for each box at level l is p_l^2 and $p_{l+1}^2 \sim p_l^2$, we can estimate the complexity of a single translation associated with level l as $O(p_l^{2\nu})$, where $\nu \geq 1$ is some parameter, that we call the “translation exponent.” For example, if the translation is performed by multiplication of the translation matrix by the vector of coefficients, the complexity of the translation will be $O(p_l^4)$, in which case $\nu = 2$.

To evaluate the number of operations required for all translations, we assume that the scatterers are distributed more or less evenly in space, and so all 8^l boxes at level l contain some scatterers. The maximum number of translations associated with any box at level l is finite (8 multipole-to-multipole translations in the upward pass, 8 local-to-local translations in the downward pass, and 189 multipole-to-local translations in the downward pass). This shows that the translation cost associated with each box at level l is $O(p_l^{2\nu})$ (with a large asymptotic constant). Therefore the total number of translations can be evaluated as

$$N_{\text{trans}} = O \left(\sum_{l=2}^{l_{\max}} p_l^{2\nu} 8^l \right). \quad (56)$$

From Eqs. (54) and (55) we can majorate the dependence of p_l on l with function $p_l = \beta k D_0 2^{-l}$, where β is some constant. Substituting this into Eq. (56), we obtain

$$N_{\text{trans}} = (k D_0)^{2\nu} O \left(\sum_{l=2}^{l_{\max}} 2^{(3-2\nu)l} \right). \quad (57)$$

Three qualitatively different cases can be identified now. The first case, $\nu < 1.5$, corresponds to “fast translations.” In this case we have $N_{\text{trans}} \sim (k D_0)^{2\nu} 2^{(3-2\nu)l_{\max}}$. Assuming that $D_0 \sim N^{1/3} a$, so $k D_0 \sim N^{1/3} ka \sim N^{1/3} p$ and $2^{3l_{\max}} \sim N$, we can see that $N_{\text{trans}} \sim p^{2\nu} N$. So the algorithm scales as $O(p^{2\nu} N)$, with p taken for the truncation of coefficients $\{\mathbf{A}^{(q)}\}$ [see Eqs. (51) and (54)]. The second case, $\nu = 1.5$, corresponds to

the “critically fast translations,” and this value of ν can be called a “critical translation exponent.” We have then, using the same assumptions as for the first case, $N_{\text{trans}} \sim (kD_0)^3 I_{\text{max}} \sim p^3 N \log N$, or complexity $O(p^3 N \log N)$. Finally, we have the case of “slow translations,” $\nu > 1.5$. Here $N_{\text{trans}} \sim (kD_0)^{2\nu} \sim p^{2\nu} N^{2\nu/3}$. So the FMM in this case scales as $O(p^{2\nu} N^{2\nu/3})$, or superlinearly with N , $2\nu/3 > 1$. It is interesting to note that in the first case the major number of translation operations are performed at the finest level of space partitioning l_{max} , in the second case the number of operations for each level is approximately the same, while in the third case the major complexity comes from operations at the coarsest level $l=2$. So two competing effects, the reduction of the number of boxes, and an increase of the truncation numbers compensate each other in the “critical” case. To provide an algorithm that scales as $O(N \log N)$, one should use at least an $O(p^3)$ translation method.

Note that the theoretically minimum translation exponent is $\nu=1$ (one operation per one expansion coefficient). This value is also “critical” for two-dimensional problems, and for three-dimensional problems, where the scatterers are distributed over some surface. The later case also applies to the boundary element methods. For these cases to have an $O(N \log N)$ complexity FMM algorithm with a variable truncation number, one should employ the theoretically fastest translation methods.

Finally, we note that the total complexity of the method consists of the complexity of translations plus complexity of generation of function representations for each box at the finest level, the complexity of evaluation of function representations, and the final summation. In the present paper all procedures have complexity $O(p^3)$. Taking into account that such operations should be performed for each point, the total complexity of the present algorithm will be $N_{\text{oper}} = O(p^3 N) + O(p^3 N \log N) = O(p^3 N \log N)$.

In the paper of Koc and Chew (1998), the diagonal forms of the translation operators obtained by Rokhlin (1993) are used. Formally, the complexity of these translation methods can be evaluated as $O(p^2)$, while their implementation requires additional procedures, such as interpolation and antinterpolation (e.g., see Darve, 2000a; Chew *et al.*, 2001), or the filtering of spherical harmonics. These procedures can be much more expensive than the translation itself [e.g., K -point Lagrange interpolation, the complexity can be evaluated as Kp^2 , and for $K \sim p$ this results in $O(p^3)$ complexity], while they can be efficient for higher frequencies. In any case, if we assume that the complexity of the translation method is $O(p^2)$ ($\nu=1$) we obtain $N_{\text{trans}} \sim O(p^2 N)$. However, higher complexity comes from the initial and final steps of the algorithm. To use the diagonal forms we need to convert the representation based on expansion coefficients to that based on samples of the signature function over the sphere. A bandlimited function with p^2 coefficients can be represented by $O(p^2)$ samples, and the conversion of one representation to the other using a direct integral evaluation (as did Koc and Chew) requires $O(p^4)$ operations. The same relates to the final evaluation step, where we should convert the function representation from the samples of the signature function to the expansion coefficients. So the total complex-

ity of the Koc and Chew algorithm can be evaluated as $O(p^4 N) + O(p^2 N) = O(p^4 N)$. There may be some possibilities to improve the complexity of the Koc and Chew algorithm, e.g., if a fast Legendre transform were available (an item of current research). We note that formally the complexities of the Koc and Chew method and the present method are the same for $p \sim \log N$, while, in practice, the quality of implementation, which influences the asymptotic constants, can reduce the execution time of the algorithm several times and factors of type $\log N$ in the range of N considered in this paper are unimportant, except from a theoretical point of view.

F. Translation method

For the present class of problems $O(p^3)$ -translation methods based on the use of the re-expansion matrices can be employed, as they provide a transparent link to the coefficients of the expansions.

The translation method based on a direct computation of the re-expansion matrices $(\mathbf{E}|\mathbf{F})(\mathbf{t})$, where $(\mathbf{E}|\mathbf{F})$ can be any combination of letters \mathbf{S} and \mathbf{R} , followed by their multiplication by the vectors of expansion coefficients has complexity $O(p^4)$. This is the lowest complexity that one can expect, even using recursive procedures for the computation of matrix elements, since the matrix is dense. To achieve $O(p^3)$ complexity, we used a translation method based on the rotation-coaxial translation decomposition of an arbitrary translation operators:

$$(\mathbf{E}|\mathbf{F})(\mathbf{t}) = \mathbf{Rot}^{-1}\left(\frac{\mathbf{t}}{t}\right) (\underline{\mathbf{E}|\mathbf{F}})(t) \mathbf{Rot}\left(\frac{\mathbf{t}}{t}\right),$$

$$t = |\mathbf{t}|, \quad \mathbf{E}, \mathbf{F} = \mathbf{S}, \mathbf{R}, \quad (58)$$

where $\mathbf{Rot}(\mathbf{t}/t)$ is the rotation transform matrix, which enables one to obtain the expansion coefficients in the reference frame rotated by such way that the z axis of the new reference frame is directed as the unit vector \mathbf{t}/t , $(\mathbf{E}|\mathbf{F})(t)$ denotes the coaxial translation matrix that performs $\overline{\text{translation}}$ along the z axis for distance t and $\mathbf{Rot}^{-1}(\mathbf{t}/t)$ is the backward rotation, which brings the reference frame to its initial orientation. Each of these operations can be performed with complexity $O(p^3)$ using a recursive computation of the matrix elements (Gumerov and Duraiswami, 2003). So, e.g., for the $(\mathbf{S}|\mathbf{R})$ translation of some vector \mathbf{C} , we first produce $\tilde{\mathbf{C}} = \mathbf{Rot}(\mathbf{t}/t)\mathbf{C}$, then $\mathbf{C}' = (\mathbf{S}|\mathbf{R})(t)\tilde{\mathbf{C}}$, and finally $\hat{\mathbf{C}} = \mathbf{Rot}^{-1}(\mathbf{t}/t)\mathbf{C}'$ for the expense of $O(p^3)$ operations, opposed to $O(p^4)$ operations if we use straightforward multiplication $\hat{\mathbf{C}} = (\mathbf{S}|\mathbf{R})(\mathbf{t})\mathbf{C}$.

1. Rotation transform

In general, an arbitrary rotation transform can be specified by three angles of rotation, e.g., the Euler angles, α_E , β_E , γ_E , or angles α , β , γ related to the Euler angles as $\alpha = \pi - \alpha_E$, $\beta = \beta_E$, and $\gamma = \gamma_E$, which are more convenient, since they are related to the spherical angles (θ_i, φ_i) of the unit vector $\mathbf{t}/t = (\sin \theta_i \cos \varphi_i, \sin \theta_i \sin \varphi_i, \cos \theta_i)$ as $\beta = \theta_i$ and $\alpha = \varphi_i$ (see Fig. 5). The rotation transform of the expansion coefficients is a function of the rotation matrix $Q(\alpha, \beta, \gamma)$. The inverse rotation transform is the same func-

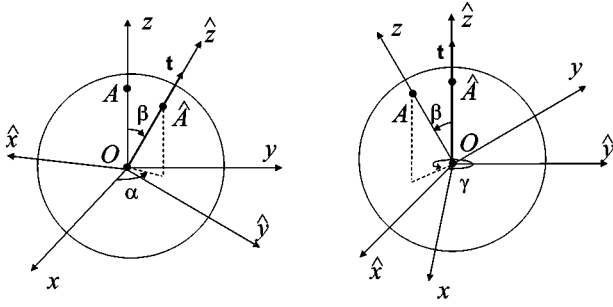


FIG. 5. An illustration of rotation specified by angles α , β , and γ . Angles β and α are the spherical polar angles of the translation vector \mathbf{t} .

tion from the inverse rotation matrix $Q^{-1}(\alpha, \beta, \gamma) = Q^T(\alpha, \beta, \gamma) = Q(\gamma, \beta, \alpha)$, so we can describe only the forward rotation, since the inverse rotation can be obtained simply by exchanging angles α and γ .

Rotation of the truncated vector \mathbf{C} with components $\{C_n^{m'}\}$ can be performed according to the following formula:

$$\tilde{C}_n^{m'} = e^{-im'\gamma} \sum_{m=-n}^n H_n^{m'm}(\beta) e^{im\alpha} C_n^m, \quad n=0,1,\dots,p-1, \quad m' = -n, \dots, n. \quad (59)$$

Here $H_n^{m'm}(\beta)$ are the entries of a real dense matrix that can be computed recursively (Gumerov and Duraiswami, 2003):

$$H_{n-1}^{m',m+1} = \frac{1}{b_n^m} \left\{ \frac{1}{2} [b_n^{-m'-1} (1 - \cos \beta) H_n^{m'+1,m} - b_n^{m'-1} (1 + \cos \beta) H_n^{m'-1,m}] - a_{n-1}^{m'} \sin \beta H_n^{m'm} \right\}, \quad n=2,3,\dots, \quad m' = -n+1, \dots, n-1, \quad m=0, \dots, n-2, \quad (60)$$

starting with the initial values

$$H_n^{m'0}(\beta) = (-1)^{m'} \sqrt{\frac{(n-|m'|)!}{(n+|m'|)!}} P_n^{|m'|}(\cos \beta), \quad n=0,1,\dots, \quad m' = -n, \dots, n. \quad (61)$$

The coefficients a_n^m and b_n^m of the recurrence relation (60) are

$$a_n^m = \begin{cases} \sqrt{\frac{(n+1+|m|)(n+1-|m|)}{(2n+1)(2n+3)}}, & n \geq |m|, \\ 0, & |m| > n, \end{cases} \quad (62)$$

$$b_n^m = \begin{cases} \sqrt{\frac{(n-m-1)(n-m)}{(2n-1)(2n+1)}}, & 0 \leq m \leq n, \\ -\sqrt{\frac{(n-m-1)(n-m)}{(2n-1)(2n+1)}}, & -n \leq m < 0, \\ 0, & |m| > n. \end{cases} \quad (63)$$

The amount of the recursive computations can be reduced using the symmetry relations:

$$H_n^{mm'}(\beta) = H_n^{m'm}(\beta) = H_n^{-m',-m}(\beta), \quad n=0,1,\dots, \quad m', m = -n, \dots, n. \quad (64)$$

So computations of the general rotation can be organized as follows. First we compute all rotation coefficients $H_n^{m'm}(\beta)$. This requires $O(p^3)$ operations. Then we perform the multiplication $D_n^m = e^{im\alpha} C_n^m$, $n=0,1,\dots,p-1$, $m=-n,\dots,n$. This can be treated as multiplication of vector \mathbf{C} by a diagonal matrix and requires $O(p^2)$ operations. Then for each subspace of degree n we perform multiplication of the dense $(2n+1) \times (2n+1)$ matrix with the entries $H_n^{m'm}(\beta)$ by vector \mathbf{D} , and this totals in $O(p^3)$ operations for all subspaces and produces the vector \mathbf{G} . The last step is multiplication of this vector by the diagonal matrix, to produce the final result: $\tilde{C}_n^{m'} = e^{-im'\gamma} G_n^{m'}$, $n=0,1,\dots,p-1$, $m' = -n, \dots, n$. This takes $O(p^2)$ operations.

We note that the decomposition can be simplified by setting $\gamma=0$, since the rotations with angles α and β align the new z axis with direction \mathbf{t}/t , while the latter rotation does not change this direction (see Fig. 5). So the forward transform can be performed with the rotation matrix $Q(\varphi_t, \theta_t, 0)$ ($\alpha = \varphi_t$, $\beta = \theta_t$, $\gamma = 0$) and the inverse with $Q(0, \theta_t, \varphi_t)$.

2. Coaxial translation

Coaxial translations, or translations along the z axis, can also be performed with lower complexity than the general translation, since each subspace of order m is invariant with respect to this type of translation. So if $\hat{\mathbf{C}} = (\mathbf{E}|\mathbf{F})(t)\mathbf{C}$, we have

$$\hat{C}_n^m = \sum_{n'=|m|}^{p-1} (E|F)_{n'n}^m(t) C_n^m, \quad m=0, \pm 1, \dots, \pm(\min(p, p')-1), \quad n'=0, 1, \dots, p'-1, \quad E, F = S, R. \quad (65)$$

Here $(E|F)_{n'n}^m$ are the entries of matrix $(\mathbf{E}|\mathbf{F})(t)$, which is truncated rectangularly, since it applies to vector \mathbf{C} of total length p^2 and produces a vector $\hat{\mathbf{C}}$ of total length $(p')^2$. This is due to the algorithm, where the truncation number must change from level to level. Since the truncation numbers p' and p are of the same order, the total complexity of the matrix-vector multiplication (65) is $O(p^3)$ for a given translation matrix. All entries of this matrix can be computed recursively for an expense of $O(p^3)$ operations using the following recursions (Gumerov and Duraiswami, 2003):

$$\begin{aligned} \text{To advance with respect to the order } n \text{ at fixed } m \geq 0: \\ a_n^m (E|F)_{n',n+1}^m = a_{n-1}^m (E|F)_{n',n-1}^m - a_{n-1}^{m'} (E|F)_{n'+1,n}^m \\ + a_{n-1}^{m'-1} (E|F)_{n'-1,n}^m, \end{aligned} \quad (66)$$

with the a 's given by Eq. (62). For advancement with respect to m :

$$\begin{aligned} b_{m+1}^{-m-1} (E|F)_{n',m+1}^m = b_n^{-m-1} (E|F)_{n'-1,m}^m \\ - b_{n+1}^m (E|F)_{n'+1,m}^m, \end{aligned} \quad (67)$$

$$n' = m + 1, m + 2, \dots, \quad E, F = S, R,$$

with the b 's given by Eq. (63). The recursive procedure starts with the initial values

$$\begin{aligned} (S|R)_{n'0}^0(t) &= (-1)^{n'} \sqrt{2n'+1} h_{n'}(kt), \\ (R|R)_{n'0}^0(t) &= (-1)^{n'} \sqrt{2n'+1} j_{n'}(kt). \end{aligned} \quad (68)$$

For the Helmholtz equation we have $(S|S)_{n'n}^m(t) = (R|R)_{n'n}^m(t)$, and so only the latter coefficients need be computed. We also note that computational savings can be achieved using the following symmetries of the coaxial translation coefficients:

$$\begin{aligned} (E|F)_{n'n}^m &= (E|F)_{n'n}^{-m} = (E|F)_{n'n}^{|m|} = (-1)^{n+n'} (E|F)_{nn'}^m, \\ E, F &= S, R. \end{aligned} \quad (69)$$

G. Use of the FMM for preconditioning in the GMRES

Our computations show that the plain (without preconditioning) GMRES slows down substantially at larger N and for relatively high volume fractions of the spheres (in this case the number of iterations increases proportionally to N). If $N_{\text{iter}} = O(N)$ then a method that requires $O(N \log N)$ per matrix-vector multiplication will require $O(N^2 \log N)$ operations to solve the problem. We found that this drawback can be fixed in many cases by the use of properly designed preconditioners. The GMRES with the right preconditioner \mathbf{M}^{-1} solves the problem

$$\mathbf{L}\mathbf{M}^{-1}(\mathbf{M}\mathbf{A}) = \mathbf{E}, \quad (70)$$

which is formally equivalent to the original problem (23) for arbitrary invertible matrix \mathbf{M} . If solution of the system $\mathbf{M}\mathbf{A} = \mathbf{F}$ for an arbitrary input \mathbf{F} can be obtained faster than solution of the original system $\mathbf{L}\mathbf{A} = \mathbf{F}$, and \mathbf{M} is a good approximation of \mathbf{L} ($\mathbf{M} \approx \mathbf{L}$), then one can expect a substantial reduction in the number of iterations.

The FMM provides a good tool for building a preconditioner (for the discussion of preconditioning see Kelly, 1995; Saad, 2003). The preconditioners that can be constructed are based on decompositions of the dense $(S|R)$ matrix in Eq. (28):

$$\begin{aligned} (S|R) &= (S|R)_{\text{near}} + (S|R)_{\text{far}}, \quad \mathbf{M}_{\text{sparse}} = \mathbf{T}^{-1} - (S|R)_{\text{near}}, \\ \mathbf{M}_{\text{dense}} &= \mathbf{T}^{-1} - (S|R)_{\text{far}}, \end{aligned} \quad (71)$$

where $(S|R)_{\text{near}}$ is a sparse matrix, which includes only interactions between the scatterers located in the same neighborhood (say at the maximum level of space subdivision), and $(S|R)_{\text{far}}$ is the matrix, which includes farfield interactions. This decomposition, in fact, is decomposition (35) in terms of the resulting vector \mathbf{B} . The matrix-vector products $\mathbf{M}_{\text{sparse}}\mathbf{A}$ or $\mathbf{M}_{\text{dense}}\mathbf{A}$ involving preconditioners $\mathbf{M}_{\text{sparse}}$ and $\mathbf{M}_{\text{dense}}$ then can be computed faster than $\mathbf{L}\mathbf{A}$, since only part of the FMM operations are needed to perform the computations. At first look, it seems that it is more reasonable to use $\mathbf{M}_{\text{sparse}}$ as the preconditioner, since presumably matrix-vector multiplication in this case should be cheaper. However, there are two reasons, both working in the same direction, why the use of $\mathbf{M}_{\text{dense}}$ may make more sense.

The first reason is based on the efficiency of the preconditioner in terms of the process convergence. Despite the fact that the distance between the neighbor scatterers is smaller and the interaction between any pair of neighbors is stronger than between the remote particles, the number of neighbors is also substantially smaller. Since this interaction decays as r^{-1} , while the number of particles in the farfield increases as r^3 (for a uniform distribution) the effect of the farfield on the given scatterer dominates for larger domain sizes and a larger number of scatterers. Most computational work then is to iterate the solution for a proper determination of the farfield, and that is what the unpreconditioned method does.

The second reason is a feature of the FMM that appears to have been used for the first time here. It is based on the efficiency of the preconditioner in terms of the speed of matrix-vector multiplication. It may seem paradoxical but in the FMM as described above the computation of $(S|R)_{\text{far}}\mathbf{A}$ can be much faster than the computation of $(S|R)_{\text{near}}\mathbf{A}$. This happens due to limitations of type (31) on the maximum level of space subdivision and takes place for a relatively high volume fraction of scatterers [for low volume fractions $(S|R)_{\text{near}}\mathbf{A}$ can be computed faster]. Indeed if l_{max} is small enough, then the most computational work is spent for direct computations of matrix-vector products $\mathbf{B}_{\text{near}} = (S|R)_{\text{near}}\mathbf{A}$ [for $l_{\text{max}} = 1$ we have $(S|R)_{\text{near}} = (S|R)$ and need to multiply the matrix by vector directly for cost $O(N^2)$]. An increase of l_{max} reduces the number of direct matrix-vector products, while increasing the total number of translations. There exist some optimum $l_{\text{max}}^{\text{opt}}$, where the costs of translations and direct computations is balanced. It appears that for high volume fractions of scatterers $l_{\text{max}}^{\text{opt}} > l_{\text{max}}$, where l_{max} is limited by condition (31) (we checked this numerically for some spatially uniform random distributions of the scatterers), and so the use of $\mathbf{M}_{\text{dense}}$ is preferable.

Finally, we note that we computed the matrix-vector product $\mathbf{A} = \mathbf{M}^{-1}\mathbf{F}$ by solving $\mathbf{M}\mathbf{A} = \mathbf{F}$ using the unpreconditioned GMRES, which introduced an internal iteration loop. As a convergence criterion for this internal process we imposed the error of iteration much larger than the error for an external loop (this is possible, since the preconditioner can be computed roughly). This type of iterative scheme produces poor results if a regular GMRES is used, due to the nonlinear dependence of \mathbf{A} on \mathbf{F} in this process. It is also well known that this drawback can be fixed by the use of flexible GMRES, which we do. For details refer to Saad (1993).

VI. NUMERICAL RESULTS AND DISCUSSION

A. Test problems computed

The multiple scattering problem even with spherical scatterers is a multiparametric problem with many degrees of freedom. In the tests we selected a few typical types of spatial distributions: uniformly random, periodic (spheres in a grid), randomly periodic, where some random pattern is repeated several times in a cubic mesh, and uniformly random with some limitations on the interparticle distances (e.g., that the ratio of the minimum distance between the sphere centers to the sphere radius, d_{min}/a , is some prescribed number). The

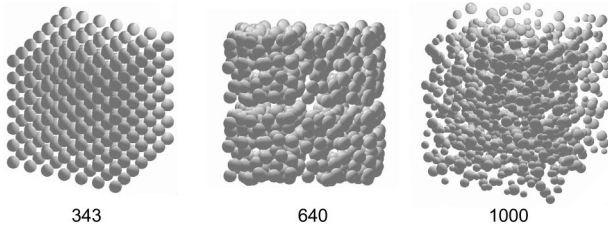


FIG. 6. Examples of distributions of N spheres used in computations: periodic, randomly periodic, and random.

latter type of distribution enables computations with high accuracy (which depends on d_{\min}/a , as discussed above). We also performed computations for monodisperse (all spheres of the same size) and polydisperse systems (see Fig. 6). In the latter case the radii of the spheres were distributed between some a_{\min} and a_{\max} (we used uniformly random distributions, while other situations are possible, depending on a particular physical problem). The number of spheres was varied in range $N=[10,10000]$. This also depends on ka , which was varied in range $0.1 < ka < 5$, and the volume fraction of the spheres in the computational domain, α :

$$\alpha = \frac{4\pi Na^3}{3D_0^3}, \quad a^3 = \sum_{q=1}^N a_q^3. \quad (72)$$

For larger ka and smaller α we used smaller number N to stay within the range $kD_0 < 100$, according to the equation

$$kD_0 = \left(\frac{4\pi N}{3\alpha}\right)^{1/3} ka. \quad (73)$$

Potentially larger kD_0 's are possible; however, some loss of precision was observed at larger kD_0 , which may be related to the number of bits representing float numbers and the growth of the cut-off errors.

We also varied the boundary admittances σ_q of the scatterers. Despite the fact that the procedure allows computations with different σ_q 's for each scatterer, we did not conduct systematic tests of this case and in all computations we assumed that all admittances are the same $\sigma_q = \sigma$, $q = 1, \dots, N$.

B. Field visualization

Once computations of the expansion coefficients $\{\mathbf{A}^{(q)}\}$ are performed, the problem is how to determine the surface and spatial distributions of the potential and the surface distributions of its normal derivative. A computation of the surface distributions is a relatively easy task, since we have (GD02),

$$\psi|_{S_q} = \frac{1}{ika_q^2} \sum_{n=0}^{p-1} \sum_{m=-n}^n \frac{A_n^{(q)m} Y_n^m(\theta_q, \varphi_q)}{kj'_n(ka_q) + i\sigma_q j_n(ka_q)}, \quad (74)$$

$$\begin{aligned} \frac{\partial \psi}{\partial n}|_{S_q} &= -\frac{\sigma_q}{ka_q^2} \sum_{n=0}^{p-1} \sum_{m=-n}^n \frac{A_n^{(q)m} Y_n^m(\theta_q, \varphi_q)}{kj'_n(ka_q) + i\sigma_q j_n(ka_q)} \\ &= -i\sigma_q \psi|_{S_q}. \end{aligned} \quad (75)$$

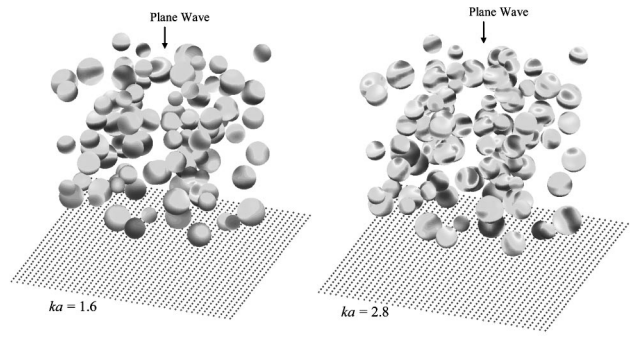


FIG. 7. Examples of the computation of the surface potential for 100 spheres exposed to a plane wave directed as shown by the arrows.

An example of surface potential visualization for the case of 100 spheres is shown in Fig. 7.

A computation of the potential at M spatial locations \mathbf{y}_m , $m=1, \dots, M$ can be a computationally more expensive task than the solution of the problem itself, if this is performed in a straightforward way. The spatial visualization, however, can be performed as fast as the solution and with the same accuracy, if the FMM is applied to this task. Two major modifications of the algorithm are needed in this case. First, we separate the set of N scatterer centers from the set of M evaluation points. The octree-based space partitioning is applied to the new data structure. As soon as the expansion coefficients $\mathbf{A}^{(q)}$, $q=1, \dots, N$ are known, the upward pass then is the same as for solution of the scattering problem, while in the downward pass the process goes over the hierarchy of the evaluation points [see the separation of data hierarchies discussed in Gumerov *et al.*, (2003)]. At the end of the downward pass we obtain the expansion coefficients $\mathbf{D}^{(n, l_{\max})}$ for each target box (the box containing the evaluation points). Then, as the final summation step, we compute the potential for all evaluation points as

$$\begin{aligned} \psi(\mathbf{y}_m) &= \psi_{\text{in}}(\mathbf{y}_m) + \mathbf{D}^{(n, l_{\max})} \cdot \mathbf{R}(\mathbf{y}_m - \mathbf{r}_*^{(n, l_{\max})}) \\ &\quad + \sum_{\mathbf{r}'_q \in E_2(n, l_{\max})} \mathbf{A}^{(q)} \cdot \mathbf{S}(\mathbf{y}_m - \mathbf{r}'_q), \\ \forall \mathbf{y}_m \in E_1(n, l_{\max}), \quad m &= 1, \dots, M. \end{aligned} \quad (76)$$

Here the first term on the right-hand side is the known potential of the incident field at the receiver locations, the second term is the potential of the field scattered by the scatterers located outside the neighborhood of the target box, and the third term is a field scattered by the scatterers located in the neighborhood of the receiver and found by direct evaluation. The complexity of this procedure for the $O(p^3)$ translation method is of type $O((N+M)\log N)$ if the same l_{\max} is used for the space imaging method as for the solution of the problem. If there are many imaging points in a target box, finer space partitioning can be used to reduce the complexity, and the optimum number of levels can be found experimentally or based on some theory applicable for the FMM (Gumerov *et al.*, 2003).

Figure 8 illustrates the “speckle patterns” on the image plane shown in Fig. 7 by blue dots. The solution is obtained using the FMM both for an accelerated iterative solution of

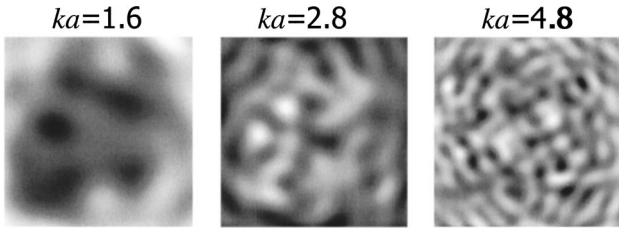


FIG. 8. Speckle patterns on the imaging plane computed using the FMM for scattering of plane waves of different wavelengths from 100 scatterers (the geometry and the imaging plane are shown in Fig. 7).

the linear system and for imaging of the field as described. The images are computed for monochromatic fields of different frequencies. These pictures are frames of a movie that animates the wave propagation in time, by plotting $\text{Re}\{\psi(\mathbf{r})e^{-i\omega t}\}$ for different t . A similar figure (Fig. 9) illustrates the “speckle patterns” for the case of 10 000 equal spheres ($a=1$) with centers that are randomly distributed inside a cube $80 \times 80 \times 80$ ($d_{\min}=2.5$).

C. A posteriori error evaluation

For these tests we used Eq. (51) to obtain the truncation number where we varied p_0 , and Eqs. (54) and (55) for the FMM procedure, where ϵ was specified to be the same as for the convergence of the iterative process. After the iterative process (GMRES) was terminated, we computed the potential at M sampling points according to Eq. (74) and the relative errors in boundary conditions using (53). The spheres were distributed randomly inside a cubic box with the restriction that the distance between the sphere centers exceeds some prescribed d_{\min} . Test were made for the incident wave $\psi_{\text{in}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}$.

An example of the computed surface error is shown in Fig. 10. This figure illustrates the behavior of $|\psi|_{S_q}(\mathbf{y}_m)$ and $\Delta_{bc}(\mathbf{y}_m)$ for $M=3200$ sampling points allocated on the surfaces of 100 equal size sound-hard spheres for $p=11$ ($ka=1$, $kD_0 \approx 30$, $\sigma=0$, $\alpha=0.042$, $d_{\min}/a=3$, $l_{\max}=3$, $\epsilon=10^{-10}$). It is seen that the maximum absolute error of computations in this case is less than 10^{-4} , while the average error of computations is of the order of 10^{-6} .

The parameter d_{\min}/a has a substantial effect on the accuracy. Figure 11 illustrates this effect for the same settings as for 10, where the truncation number and d_{\min}/a were varied. The relative error $\epsilon^{(bc)}$ defined by Eq. (53), that also can

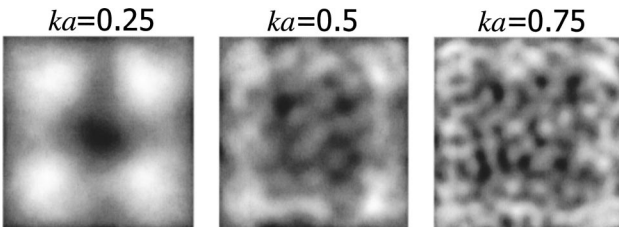


FIG. 9. Speckle patterns on the imaging plane computed using the FMM for scattering of plane waves of different wavelength from 10 000 spheres of size $a=1$ randomly distributed inside a cube $80 \times 80 \times 80$. The image plane is located at $z=82$ (assuming that the computational domain is between $z=-1$ and $z=81$ and the plane wave propagates in the positive direction).

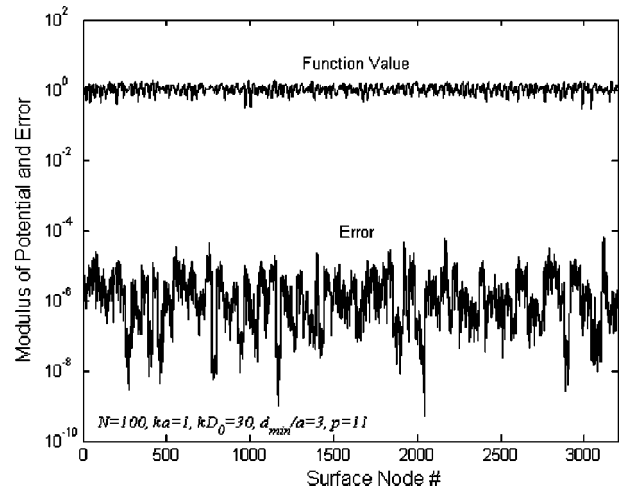


FIG. 10. The modulus of the potential and the absolute error in boundary conditions obtained *a posteriori* as functions on the surface node index. The surface of each of 100 spheres was sampled by 32 points.

be assigned as the relative error in the “infinity” norm L_∞ , is much larger (about one order of magnitude) than the average, or L_2 -norm error, $\epsilon_2^{(bc)}$, defined as

$$\epsilon_2^{(bc)} = \frac{1}{M} \left\{ \sum_{m=1}^M [\epsilon_m^{(bc)}]^2 \right\}^{1/2}, \quad (77)$$

with $\epsilon_m^{(bc)}$ specified in Eq. (53). It is seen that even for the case when the spheres may touch ($d_{\min}/a=2$), computations can be performed with relatively small truncation numbers if an accuracy $\epsilon^{(bc)}$ of the order of a few percent is acceptable (that might be sufficient for comparisons with experiments, field imaging, etc.). For computations with high precision, the truncation numbers should be increased, which may substantially slow down the computation process. One of the methods to treat this drawback is to introduce truncation numbers p_q for each scatterer, $q=1, \dots, N$, which should depend on the distance from the q th sphere to its closest neighbor sphere. In the present study, however, we did not perform

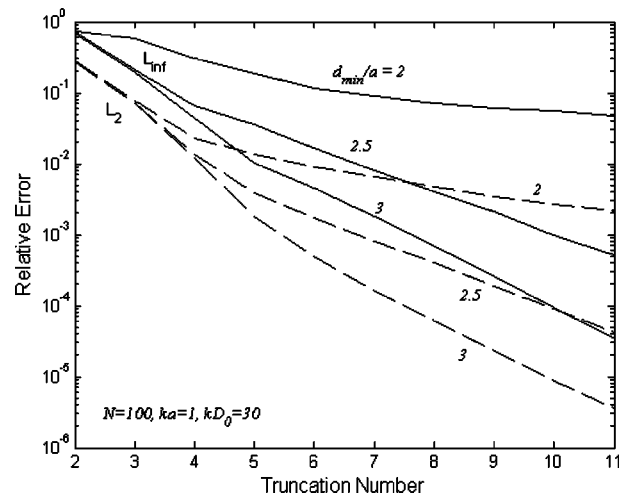


FIG. 11. Dependencies of the relative errors in the infinite (solid lines) and quadratic (dashed lines) norms on the truncation number p for the same parameters as shown in Fig. 10. The errors are computed for random distributions with different ratios d_{\min}/a shown near the curves.

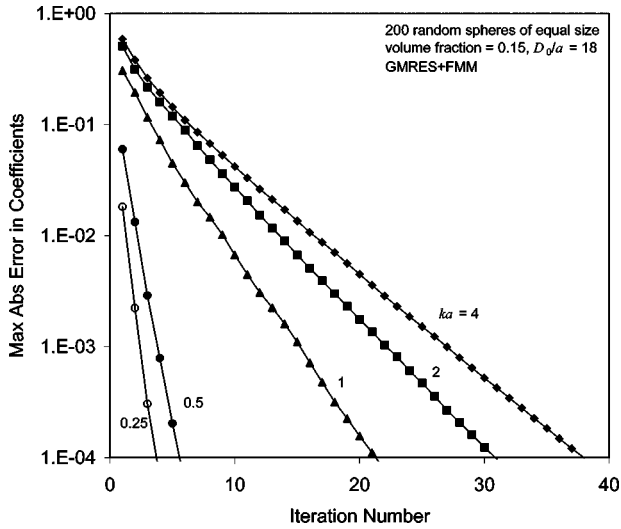


FIG. 12. Convergence of the unpreconditioned GMRES at different dimensionless wave numbers ka shown near the curves.

modifications, since our goal was to develop and test the base algorithm.

D. Convergence

Several factors may affect the rate of convergence of the iterative method and more or less a complete study involves the investigation of the multiparametric space. In the present tests we usually fixed all parameters for the case and varied one or two parameters to see their effect on the convergence rate and accuracy of computations. Figure 12 shows the effect of the frequency on the convergence of unpreconditioned GMRES. For the case illustrated, 200 equal size sound-hard spheres with random distribution of their centers were generated as described above ($d_{\min}/a=2.3$). The truncation numbers were selected according to Eq. (51) with a constant $p_0=10$ for all cases, while the FMM truncation numbers were determined according to Eq. (54) with $\epsilon=10^{-4}$. It is seen that in this case the number of iterations to achieve specified accuracy increases with ka . An *a posteriori* error check for the boundary conditions at 6400 points shows that for this case the error $\epsilon^{(bc)}$ defined by Eq. (53) does not exceed 5×10^{-3} for all cases, while $\epsilon_2^{(bc)}$ defined by Eq. (77) is about 1 order of magnitude smaller. We found that the iteration process for higher frequencies ($ka \geq 5$, $kD_0 \geq 100$) diverges, and special methods for treatment of higher frequencies are required. Several methods can be tried in this case, such as using preconditioners based on high-frequency asymptotic expansions of the system matrix, and some relaxation methods, where, e.g., the impedance of the spheres, σ , can be gradually adjusted during the iteration process. Indeed, we did some preliminary computations with fixed $ka=4$ and varying σ and found that, say for $\sigma/k \sim 0.1$, the number of iterations to achieve the same accuracy drops by a factor of 2.

Properly designed preconditioners can substantially reduce the computation time to obtain a solution of specified accuracy. Some results of the use of GMRES and FGMRES with the right dense preconditioner $\mathbf{M}_{\text{dense}}$ [see Eq. (71)] is shown in Table I. This table shows computational results for

TABLE I. Performance of the FMM accelerated GMRES/FGMRES. $N=10\,000$. Number of complex unknowns= 2.25×10^6 .

Preconditioner	No. external iterations	No. internal iterations	CPU time (s)
No preconditioner	58	0	8340
Right dense, $\epsilon_{\text{pre}}=0.1$	11	87	3460
Right dense, $\epsilon_{\text{pre}}=0.01$	10	165	4525

10 000 sound hard spheres of equal size with a randomly periodic distribution (blocks of 80 spheres with random distribution of centers, $d_{\min}/a=2$, $\alpha=0.2$, were repeated 125 times to occupy a larger cube). Here $D_0/a=60$ and $ka=0.5$, $l_{\max}=4$, $\epsilon=10^{-4}$, $p=7$, and the CPU time was measured for 3.2 MHz dual Intel Xeon processor (3.5 GB RAM, 25% CPU resource allocation for a task). The table shows that the number of iterations for the external loop can be drastically reduced due to the use of a preconditioner. The price to pay for this is the internal iterative loop, which uses also GMRES to obtain a solution of the equation $\mathbf{M}_{\text{dense}}\mathbf{A}=\mathbf{F}$. However, for the present case matrix-vector multiplication $\mathbf{M}_{\text{dense}}\mathbf{A}$ was seven times faster than performing multiplication $\mathbf{L}\mathbf{A}$ (both using the FMM). This resulted in a decrease of the overall computational time. A solution of equation $\mathbf{M}_{\text{dense}}\mathbf{A}=\mathbf{F}$ can be performed with much less accuracy, ϵ_{pre} (shown in the first column of the table), than the accuracy required for the external iterative process. This can be a subject for optimization, since the increase of the accuracy for internal iterations reduces the number of external iterations and increases the number of the internal iterations. We also need to mention that the efficiency of preconditioners depend on several other parameters, including wave number, number of scatterers, surface admittance, volume fraction, etc. (See Table I)

E. Performance study

To determine the dependence of the CPU time required for a solution of the problem, and to check the consistency of the actual results with the theoretical predictions, we conducted a performance study. All computations were made in double precision complex arithmetic on a desktop PC (3.2 MHz dual Intel Xeon processor, 3.5 GB RAM, 25% CPU resource allocated for a task). As a test distribution we used randomly periodic distributions of equal size sound hard spheres (blocks of 80 spheres with random distribution of centers, $d_{\min}/a=2$, were repeated m^3 times to occupy a larger cube, $m=1, \dots, 5$; an example for $m=2$, $N=640$ is shown in Fig. 6). Since the volume fraction and the wave number in the tests were fixed ($\alpha=0.2$, $ka=0.5$) the increase in the number of spheres also means the increase of the size of the computational domain [see Eq. (73)].

We found that to reduce the CPU time for the present case, one should select the computational domain size larger than the size of the box, which encloses all the spheres. The size of the computational domain can be determined based on l_{\max} according Eq. (33) as

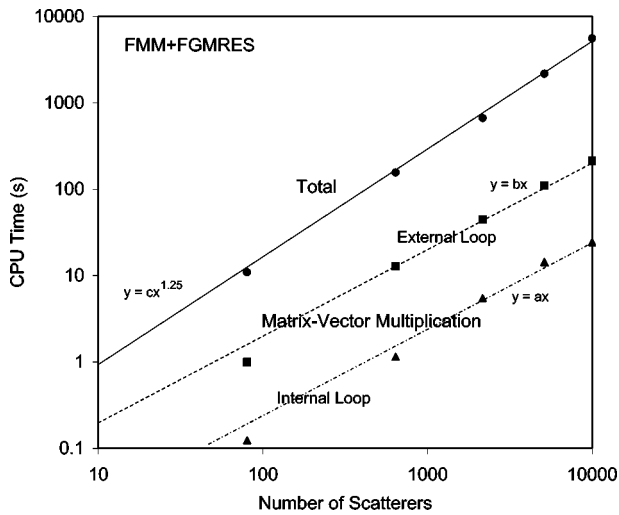


FIG. 13. Dependences of the CPU time (Intel Xeon 3.2 GHz, 3.5 Gb RAM) on the number of scatterers ($ka=0.5$, $p=15$, $\epsilon=10^{-4}$) for randomly-periodic distribution of sound-hard spheres of equal size (see Fig. 6). The triangles and squares show the CPU time required for matrix-vector multiplication using the FMM for the preconditioner and full matrix, respectively. The disks show the total CPU time for the execution of the algorithm. The lines show the interpolation of data.

$$D_0 = \frac{2\delta_{\min}}{4 - (1 + \delta_{\min})\sqrt{3}} 2^{l_{\max}} a_{\max}. \quad (78)$$

Setting here $l_{\max}=2,3,\dots$, we can determine a discrete set of D_0 . For computations we selected the minimum of D_0 from this set, which encloses all the spheres. Although δ_{\min} here should be larger than 1, we found with the *a posteriori* error check that $\delta_{\min}=1$ provides stable and accurate results in this case. Note that for performance tests we accepted relative errors $\epsilon^{(bc)}$ of about 2% (the errors $\epsilon_2^{(bc)}$ were one order of magnitude smaller). To provide this order of errors we set $\epsilon=10^{-4}$ for the global iteration convergence, while the error for the preconditioner was set to $\epsilon_{\text{pre}}=0.2$. In all computations we used a truncation number of $p=15$. This substantially exceeds $p_{l_{\max}} \sim 8$, and the choice of p was dictated by a desire to compute accurately the case of touching spheres. Since the excessive value $p=15$ is due only to the influence of the neighbor spheres, for each sphere we took the contribution of the neighbors with $d/a < 1.3$ into harmonics of degree $n=8,\dots,15$, while we neglected the effect of other scatterers on these harmonics (see the discussion at the end of Sec. V D). An *a posteriori* error check showed that the relative errors specified above were achieved.

Figure 13 shows the measurements of the CPU time for various operations of the algorithm based on the FGMRES with a right preconditioner. It is seen that the computation time scales almost linearly for a single matrix-vector multiplication in the internal or external loop using the FMM. This agrees well with the theoretical predictions of complexity $O(N \log N)$. Due to the growth of the number of iterations both in the internal and external loops with N , the total CPU time scales as $O(N^{1.25})$. This shows that some average number of internal-external iterations, N_{iter} , is proportional in the present case to $N^{0.25}$ or so. As it was discussed previously, an optimum can be sought by varying ϵ_{pre} , and pos-

sibly, ϵ , to perform better, while we kept these parameters for the illustration case fixed. Note that for large N and unpreconditioned GMRES we found that the number of iterations N_{iter} grows proportionally to N for $N \geq 2000$, which substantially reduces the efficiency of the algorithm.

It is also noteworthy that even for $N=80$ the size of the complex system matrix \mathbf{L} is $Np^2 \times Np^2 = 18\,000 \times 18\,000$, which would require several hours to solve the problem directly, while the present method requires only 10 s for this operation. We also found that a solution of this problem using GMRES without preconditioners and the FMM for matrix-vector multiplication takes about 75 s of CPU time for the same machine. Cases with larger N were not even computable due to memory limitations.

VII. CONCLUSIONS

We have developed and tested a procedure for a solution of the acoustic scattering problems for the case of clusters consisting of a relatively large number of spherical scatterers of various radii and impedances arbitrarily located in three-dimensional space. This procedure uses a multipole re-expansion technique, the fast multipole method, and iteration algorithms, such as GMRES with possible preconditioners and is accelerated utilizing the FMM. We developed and tested a technique for *a posteriori* error control, investigated convergence, and performance of the method in a certain range of parameters and found that this technique can be an efficient, accurate, and powerful tool for solution of such multiple scattering problems.

While the method is developed and tested for spherical scatterers, it can be naturally extended for the computation of scattering from objects of an arbitrary shape, as soon as the T matrices for the single scatterers are known analytically or precomputed numerically. Several research issues for further improvement of the method remain open, e.g., the problem of computations for very high frequencies or large domain sizes, $kD_0 \geq 100$.

ACKNOWLEDGMENTS

We would like to gratefully acknowledge the support of NSF awards No. 0219681 and No. 0086075. We would also like to thank Professor Howard Elman and Professor Dianne O'Leary for discussions and guidance related to iterative methods used in the present study. We would also like to thank Professor Rajarshi Roy for suggesting the "speckle pattern" problem.

- Nigmatulin, R. I. (1990). *Dynamics of Multiphase Media* (Hemisphere, Washington, DC).
- Caflich, R., Miksis, M., Papanicolaou, G. C., and Ting, L. (1985). "Effective equations for wave propagation in bubbly liquids," *J. Fluid Mech.* **153**, 259–273.
- Gumerov, N. A., Ivandaev, A. I., and Nigmatulin, R. I. (1988). "Sound waves in monodisperse gas-particle or vapor-droplet mixtures," *J. Fluid Mech.* **193**, 53–74.
- Duraiswami, R., and Prosperetti, A. (1995). "Linear pressure waves in fogs," *J. Fluid Mech.* **299**, 187–215.
- Waterman, P. C., and Truell, R. (1961). "Multiple scattering of waves," *J. Math. Phys.* **2**, 512–537.

- Peterson, B., and Strom, A. (1974). "Matrix formulation of acoustic scattering from an arbitrary number of scatterers," *J. Acoust. Soc. Am.* **56**, 771–780.
- Varadan, V. K., and Varadan, V. V. (1980). *Acoustic, Electromagnetic and Elastic Wave Scattering: Focus on the T-Matrix Approach* (Pergamon, New York).
- Wang, Y. M., and Chew, W. C. (1993). "A recursive T-matrix approach for the solution of electromagnetic scattering by many spheres," *IEEE Trans. Antennas Propag.* **41**, 1633–1639.
- Mishchenko, M. I., Travis, L. D., and Mackowski, D. W. (1996). "T-matrix computations of light scattering by nonspherical particles: a review," *J. Quant. Spectrosc. Radiat. Transf.* **55**, 535–575.
- Koc, S., and Chew, W. C. (1998). "Calculation of acoustical scattering from a cluster of scatterers," *J. Acoust. Soc. Am.* **103**, 721–734.
- Gumerov, N. A., and Duraiswami, R. (2002). "Computation of scattering from N spheres using multipole reexpansion," *J. Acoust. Soc. Am.* **112**, 2688–2701.
- Ong, E. T., Heow, P. L., and Kian, M. L. (2004). "A fast Fourier transform on multipoles (FFTM) algorithm for solving Helmholtz equation in acoustic analysis," *J. Acoust. Soc. Am.* **116**, 1362–1371.
- Greengard, L., and Rokhlin, V. (1987). "A fast algorithm for particle simulations," *J. Comput. Phys.* **73**, 325–348.
- Sangani, A. S., and Mo, G. (1996). "An $O(N)$ algorithm for Stokes and Laplace interactions of particles," *Phys. Fluids* **8**, 1990–2010.
- Rokhlin, V. (1993). "Diagonal forms of translation operators for the Helmholtz equation in three dimensions," *Appl. Comput. Harmon. Anal.* **1**, 82–93.
- Chew, W. C., Jin J.-M., Michielssen, E., and Song, J. (2001). *Fast and Efficient Algorithms in Computational Electromagnetics* (Artech House, Boston).
- Darve, E. (2000a). "The fast multipole method: numerical implementation," *J. Comput. Phys.* **160**, 195–240.
- Gumerov, N. A., and Duraiswami, R. (2003). "Recursions for the computation of multipole translation and rotation coefficients for the 3-D Helmholtz equation," *SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput.* **25**, 1344–1381.
- Saad, Y. (2003). *Iterative Methods for Sparse Linear Systems*, 2nd ed. (SIAM, Philadelphia).
- Kelly, C. T. (1995). *Iterative Methods for Linear and Nonlinear Equations* (SIAM, Philadelphia).
- Greengard, L. (1988). *The Rapid Evaluation of Potential Fields in Particle Systems* (MIT Press, Cambridge, MA).
- Gumerov, N. A., Duraiswami, R., and Borovikov, E. A. (2003). "Data Structures, Optimal Choice of Parameters, and Complexity Results for Generalized Multilevel Fast Multipole Methods in d Dimensions," University of Maryland Institute for Advanced Computer Studies Technical Report UMIACS-TR-#2003-28; also CS-TR-#4458. (Available at <http://www.cs.umd.edu/Library/TRs/CS-TR-4458/CS-TR-4458.pdf>.)
- Koc, S., Song, J., and Chew, W. C. (1999). "Error analysis for the numerical evaluation of the diagonal forms of the scalar spherical addition theorem," *SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal.* **36**, 906–921.
- Darve, E. (2000b). "The fast multipole method I: error analysis and asymptotic complexity," *SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal.* **38**, 98–128.
- Saad, Y. (1993). "A flexible inner-outer preconditioned GMRES algorithm," *SIAM J. Sci. Comput. (USA)* **14**, 461–469.

Modeling of phased array transducers

Rais Ahmad and Tribikram Kundu^{a)}

Department of Civil Engineering and Engineering Mechanics, University of Arizona, Tucson, Arizona 85721

Dominique Placko

Ecole Normale Supérieure, SATIE, 61, av. du Président Wilson F-94235 Cachan cedex, France

(Received 4 June 2004; revised 26 October 2004; accepted 27 October 2004)

Phased array transducers are multi-element transducers, where different elements are activated with different time delays. The advantage of these transducers is that no mechanical movement of the transducer is needed to scan an object. Focusing and beam steering is obtained simply by adjusting the time delay. In this paper the DPSM (distributed point source method) is used to model the ultrasonic field generated by a phased array transducer and to study the interaction effect when two phased array transducers are placed in a homogeneous fluid. Earlier investigations modeled the acoustic field for conventional transducers where all transducer points are excited simultaneously. In this research, combining the concepts of delayed firing and the DPSM, the phased array transducers are modeled semi-analytically. In addition to the single transducer modeling the ultrasonic fields from two phased array transducers placed face to face in a fluid medium is also modeled to study the interaction effect. The importance of considering the interaction effect in multiple transducer modeling is discussed, pointing out that neighboring transducers not only act as ultrasonic wave generators but also as scatterers. © 2005 Acoustical Society of America.

[DOI: 10.1121/1.1835506]

PACS numbers: 43.20.Fn, 43.20.Bi, 43.20.Ei [JBS]

Pages: 1762–1776

I. INTRODUCTION

A phased array transducer is composed of many elements arranged in a certain pattern that emit acoustic energy at different times. The elements can be pulsed in certain sequence to control the beam angle, also known as the steering angle. During conventional ultrasonic nondestructive testing (NDT) experiments, one transducer mechanically moving from one point to the next or several transducers working simultaneously inspect an object. Mechanical movement of one transducer is a time-consuming operation and the use of multiple transducers is an expensive operation. On the other hand, one phased array transducer can scan an object relatively quickly because it can inspect a wide range by emitting acoustic energy in controlled directions. Thus efficient scanning is possible by phased-array transducers. Wooh and Shi (1999) and Azar and Wooh (1999a, b) experimentally studied the steering and focusing behavior of phased array transducers in concrete structures. In the present paper phased array transducers are semi-analytically modeled by the distributed point source method (DPSM). The DPSM is a semi-analytical technique used to model ultrasonic (or eddy current) transducers for computing pressure and velocity fields (or magnetic fields) generated by the transducers. Placko and Kundu (2001, 2003) first developed the method for ultrasonic transducer modeling. This technique has been used to model ultrasonic fields in homogeneous and nonhomogeneous fluids, and scattered fields generated by a solid scatterer of finite dimension placed in the fluid (Placko *et al.*, 2001, 2002, 2003). Inhomogeneous media composed of fluid

and solid materials with a fluid–solid interface have also been modeled by this technique (Lee *et al.*; 2002). Use of the DPSM for solving problems other than ultrasonic problems is discussed in Placko *et al.*, (2001). In the DPSM technique, the active part of the transducer surface is discretized into a finite number of small hemispherical surface areas. Point sources are placed at the centers of these hemispheres. Each point source acts as an active source of energy. For each point source, an analytical solution for the generated field can be obtained. The total ultrasonic field is computed by superimposing the solutions of all point sources.

In this research work, a phased array transducer is modeled using the DPSM to obtain the ultrasonic field in a homogeneous medium. For this purpose the transducer face is discretized and point sources are placed in every elemental area. The point sources are excited with a certain phase difference (or time lag) to enhance the dynamic focusing. Later, interaction between two phased array transducers placed in a homogeneous medium is studied. The results presented here compare the phased array transducer and conventional planar transducer generated ultrasonic beams in different directions. The results also show the importance of considering the interaction effect between two phased array transducers when the total field is computed. For computing a mathematically correct ultrasonic field, it should be kept in mind that a transducer not only acts as a generator of the ultrasonic energy but also as a scatterer of the energy generated by the other transducer. However, the scattering effect is found to be small in many cases considered in this paper.

II. THEORY

The distributed point source method can be used to compute the pressure field generated by an ultrasonic transducer

^{a)} Author to whom correspondence should be addressed. Electronic mail: tkundu@email.arizona.edu

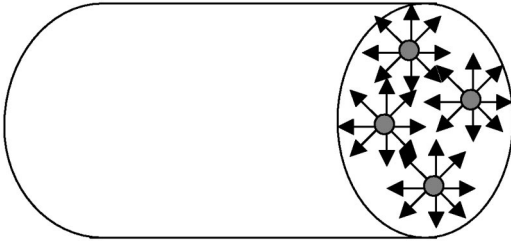


FIG. 1. Four point sources distributed over a finite surface of an ultrasonic transducer.

of finite dimension, placed in a homogeneous or nonhomogeneous fluid (or fluid–solid) media, or in a fluid with a scatterer of finite dimension. The theory associated with this formulation is briefly described in the following. For more detailed description the readers are referred to Placko and Kundu (2003) and Ahmad *et al.* (2004).

A. Planar transducer modeling by the distributed of point source method (DPSM)

The basic principle of the DPSM for modeling of magnetic and ultrasonic sensors has been described in a number of earlier publications (Placko and Kundu, 2003). The active surface of the transducer is discretized into an array of point sources. Sources on the transducer surface are shown in Fig. 1. Each point source transmits a signal. The total signal is the superposition of the signals transmitted by all point sources. The pressure field generated by a plane source of finite dimension can be assumed to be the summation of the pressure fields generated by a number of point sources distributed over the surface of the finite source as shown in Fig. 1. The finite plane source can be, for example, the front face of a transducer as shown in the figure. The combined effect of a large number of point sources distributed on a plane surface is the vibration of the particles in the direction normal to the plane surface. Non-normal components of motion at a point on the surface, generated by neighboring source points, cancel each other.

B. Computation of an ultrasonic field in a homogeneous fluid using the distributed point source method

For a flat surface transducer, where all point sources are excited at the same time, the pressure field generated at a point \mathbf{x} (see Fig. 2) by the transducer can be obtained by integrating the spherical waves, as done in the conventional surface integral technique (Schmerr, 1998; Placko and Kundu, 2003),

$$p(\mathbf{x}) = \int_S B \frac{\exp(ik_f r)}{4\pi r} dS, \quad (1)$$

where B relates the source velocity and pressure and, thus, proportional to the source velocity amplitude, k_f is the wave number and r is the distance between the observation point \mathbf{x} and the source points distributed over the transducer face.

For a phased array transducer, the point sources are excited at different times. Let the strength of the m th point

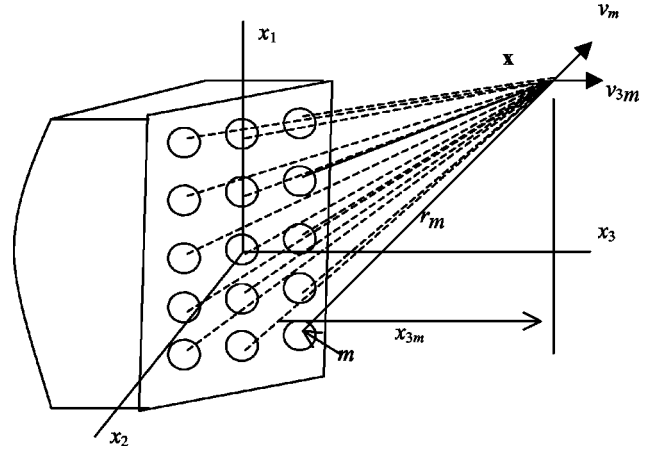


FIG. 2. Velocity v_m at point \mathbf{x} due to the m th point source.

source be A_m , and it is excited at a time t_m . Then the pressure at a distance r_m from the point source is given by

$$p_m(r) = A_m \frac{\exp(ik_f r_m - i\omega t_m)}{r_m}, \quad (2)$$

where ω is the angular frequency. For the phased array transducer, the point sources on the surface are assumed to be excited after different time intervals Δt_m ; in Eq. (2) the term t_m can be expressed in terms of a time lag from a reference time t ,

$$t_m = t - \Delta t_m. \quad (3)$$

If there are N point sources distributed over the transducer surface, as shown in Fig. 2, then the total pressure at point \mathbf{x} is given by

$$\begin{aligned} p(\mathbf{x}) &= \sum_{m=1}^N p_m(r_m) \\ &= \sum_{m=1}^N A_m \frac{\exp(ik_f r_m - i\omega(t - \Delta t_m))}{r_m}, \end{aligned} \quad (4)$$

where r_m is the distance of the m th point source from point \mathbf{x} .

It should be noted here that Eq. (1) can be written in discrete summation form in the following manner:

$$p(\mathbf{x}) = \sum_{m=1}^N B_m \frac{\exp(ik_f r_m)}{4\pi r_m} \Delta S_m. \quad (5)$$

In Eq. (5) the time dependence $e^{-i\omega t}$ is implied. If Eq. (4) is specialized for the conventional or static transducer case where all elements are excited at the same phase, then Eq. (4) is simplified to

$$p(\mathbf{x}) = \sum_{m=1}^N A_m \frac{\exp(ik_f r_m)}{r_m}. \quad (6)$$

In Eq. (6) the time dependence $e^{-i\omega t}$ is implied. Comparing Eqs. (5) and (6) the relation between A_m and B_m can be obtained,

$$A_m = \frac{B_m \Delta S_m}{4\pi}. \quad (7)$$

The particle velocity in the radial direction, generated by the m th point source at a distance r , is given by

$$v_m(r) = \frac{A_m}{i\omega\rho} \frac{\partial}{\partial r} \left(\frac{\exp(ik_f r - i\omega(t - \Delta t_m))}{r} \right)$$

$$= \frac{A_m}{i\omega\rho} \left(\frac{ik_f \exp(ik_f r)}{r} - \frac{\exp(ik_f r)}{r^2} \right)$$

$$\times \exp(-i\omega(t - \Delta t_m)),$$

or

$$v_m(r) = \frac{A_m}{i\omega\rho} \frac{\exp(ik_f r)}{r} \left(ik_f - \frac{1}{r} \right)$$

$$\times \exp(-i\omega(t - \Delta t_m)). \quad (8)$$

After contributions of all N sources are added, the velocity components in the three orthogonal directions x_1 , x_2 , and x_3 (see Fig. 2) are obtained,

$$v_1(\mathbf{x}) = \sum_{m=1}^N v_{1m}(r_m) = \sum_{m=1}^N \frac{A_m}{i\omega\rho} \frac{x_{1m} \exp(ik_f r_m)}{r_m^2}$$

$$\times \left(ik_f - \frac{1}{r_m} \right) \exp(-i\omega(t - \Delta t_m)),$$

$$v_2(\mathbf{x}) = \sum_{m=1}^N v_{2m}(r_m) = \sum_{m=1}^N \frac{A_m}{i\omega\rho} \frac{x_{2m} \exp(ik_f r_m)}{r_m^2}$$

$$\times \left(ik_f - \frac{1}{r_m} \right) \exp(-i\omega(t - \Delta t_m)), \quad (9)$$

$$v_3(\mathbf{x}) = \sum_{m=1}^N v_{3m}(r_m) = \sum_{m=1}^N \frac{A_m}{i\omega\rho} \frac{x_{3m} \exp(ik_f r_m)}{r_m^2}$$

$$\times \left(ik_f - \frac{1}{r_m} \right) \exp(-i\omega(t - \Delta t_m)).$$

On the transducer surface $v_1=0$ and $v_2=0$ because the transducer surface vibrates with a nonzero velocity in the x_3 direction only. Note that for conventional transducers $v_3=v_0$ (the time dependence $e^{-i\omega t}$ is implied) and for phased array transducers $v_3=v_0 \exp\{-i\omega(t - \Delta t_m)\}$. If the observation point \mathbf{x} is placed on the transducer surface, then the three expressions of Eq. (9) should be equal to 0, 0, and v_0 (or $v_0 \exp\{-i\omega(t - \Delta t_m)\}$), respectively. Thus, from N points on the transducer surface, $3N$ equations are obtained. However, we have only N unknowns (A_1, A_2, \dots, A_N) in the above equation set. To obtain the same number of unknowns as equations the number of unknowns can be increased from N to $3N$ by replacing each point source by a triplet. In triplet modeling the single source of every sphere is replaced by three sources. Three point sources of each triplet are located at three vertices of an equilateral triangle and these triangles are oriented randomly on a plane at $x_3 = -r_s$. Random orientations of triplet sources are necessary to preserve the iso-

tropic material properties and prevent any preferential orientation. It should be noted that the triplet source modeling is not needed if only the v_3 component of displacement is equated to v_0 .

For viscous fluid and solid interface modeling all three velocity components should be matched across an interface, but that is not necessary for ideal or perfect fluid modeling. In this paper ideal fluids are considered and only the v_3 component of velocity is matched at the transducer–fluid interface. However, it was observed that the ultrasonic field computed by matching only the v_3 component of velocity was very close to the one obtained by matching all three velocity components, if the number of point sources considered in the triplet modeling and simple (or single) source modeling were kept same. In other words, N triplet sources (consisting of $3N$ point sources) generate an ultrasonic field that is very close to the one produced by $3N$ simple point sources. Since the total number of point sources used in the DPSM modeling controls the computation time, both triplet and simple point source modeling takes almost the same amount of computer time for a similar degree of accuracy in computation.

The point sources are placed at $x_3 = -r_s$ to guarantee a nonvanishing denominator in Eq. (9). Note that if the point sources are placed on the transducer surface, then for some observation points on the transducer surface the distance (r_m) between the m th source and the observation point may vanish, giving rise to the numerical error (division by zero) while computing the velocity components using Eq. (9). For this reason point sources are placed at $x_3 = -r_s$, slightly behind the transducer face at $x_3 = 0$. The length r_s is obtained by equating the total hemi-spherical surface area ($2\pi N r_s^2$) of N hemispheres of radius r_s to the front surface area (A) of the transducer. Note that N is the number of point sources used to model the transducer face.

After solving the system of $3N$ linear equations (for triplet sources) or a system of N linear equations (for simple point sources obtained by equating only v_3 to v_0 or $v_0 \exp\{-i\omega(t - \Delta t_m)\}$) the source strengths A_m associated with all point sources are obtained. After evaluating A_m , the pressure $p(\mathbf{x})$ can be calculated at any point (on the transducer surface or away) from Eq. (4).

C. Matrix formulation

Equation (9) can be written in the form of a matrix equation,

$$\mathbf{V}_S = \mathbf{M}_{SS} \mathbf{A}_S, \quad (10)$$

where \mathbf{V}_S is the $(3N \times 1)$ vector of the velocity components at N number of surface points and \mathbf{A}_S is the $(3N \times 1)$ vector containing the strengths of $3N$ point sources. \mathbf{M}_{SS} is the $(3N \times 3N)$ matrix relating the two vectors \mathbf{V}_S and \mathbf{A}_S , where

$$\{\mathbf{V}_S\}^T = [v_1^1 \quad v_2^1 \quad v_3^1 \quad v_1^2 \quad v_2^2 \quad v_3^2 \quad \cdots \quad v_1^N \quad v_2^N \quad v_3^N], \quad (11)$$

$$\{\mathbf{A}_S\}^T = [A_1 \quad A_2 \quad A_3 \quad A_4 \quad A_5 \quad A_6 \quad \cdots \quad A_{3N-2} \quad A_{3N-1} \quad A_{3N}], \quad (12)$$

and

$$\mathbf{M}_{SS} = \begin{bmatrix} f(x_{11}^1, r_1^1) & f(x_{12}^1, r_2^1) & f(x_{13}^1, r_3^1) & \dots & \dots & f(x_{1(3N-1)}^1, r_{3N-1}^1) & f(x_{1(3N)}^1, r_{3N}^1) \\ f(x_{21}^1, r_1^1) & f(x_{22}^1, r_2^1) & f(x_{23}^1, r_3^1) & \dots & \dots & f(x_{2(3N-1)}^1, r_{3N-1}^1) & f(x_{2(3N)}^1, r_{3N}^1) \\ f(x_{31}^1, r_1^1) & f(x_{32}^1, r_2^1) & f(x_{33}^1, r_3^1) & \dots & \dots & f(x_{3(3N-1)}^1, r_{3N-1}^1) & f(x_{3(3N)}^1, r_{3N}^1) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ f(x_{31}^N, r_1^N) & f(x_{32}^N, r_2^N) & f(x_{33}^N, r_3^N) & \dots & \dots & f(x_{3(3N-1)}^N, r_{3N-1}^N) & f(x_{3(3N)}^N, r_{3N}^N) \end{bmatrix}_{3N \times 3N}, \quad (13)$$

where

$$f(x_{jm}^n, r_m^n) = \frac{x_{jm}^n \exp(ik_f r_m^n)}{i\omega\rho(r_m^n)^2} \left(ik_f - \frac{1}{r_m^n} \right) \times \exp(-i\omega(t - \Delta t_m^n)). \quad (14)$$

In Eq. (14), the first subscript j of x can take value 1, 2, or 3 and indicate whether x is measured in the x_1 , x_2 , or x_3 direction. The subscript m of x and r can take values from 1 to $3N$ depending on which point source is considered and the superscript n can take any value between 1 and N corresponding to the point on the transducer surface where the velocity component is computed. In this formulation for $3N$ point sources, three boundary conditions on the velocity are satisfied at every point of the N boundary points.

Source strength vectors can be obtained as

$$\mathbf{A}_S = [\mathbf{M}_{SS}]^{-1} \mathbf{V}_S = \mathbf{N}_{SS} \mathbf{V}_S. \quad (15)$$

The pressure \mathbf{P}_T and velocity vector \mathbf{V}_T at any given number (M) of points $\{\mathbf{x}\}$ on a target surface can be simultaneously obtained by

$$\begin{aligned} \mathbf{P}_T &= \mathbf{Q}_{TS} \mathbf{A}_S, \\ \mathbf{V}_T &= \mathbf{M}_{TS} \mathbf{A}_S. \end{aligned} \quad (16)$$

Matrix \mathbf{M}_{TS} is same as \mathbf{M}_{SS} of Eq. (13) when the target points or reference points are identical to the transducer surface points where the velocity components are matched to obtain the point source strength vector \mathbf{A}_S in Eq. (15). Matrix \mathbf{Q}_{TS} is obtained as

$$\mathbf{Q}_{TS} = \begin{bmatrix} \frac{\exp(ik_f r_1^1 - i\omega(t - \Delta t_1^1))}{r_1^1} & \frac{\exp(ik_f r_2^1 - i\omega(t - \Delta t_2^1))}{r_2^1} & \dots & \dots & \frac{\exp(ik_f r_{3N}^1 - i\omega(t - \Delta t_{3N}^1))}{r_{3N}^1} \\ \frac{\exp(ik_f r_1^2 - i\omega(t - \Delta t_1^2))}{r_1^2} & \frac{\exp(ik_f r_2^2 - i\omega(t - \Delta t_2^2))}{r_2^2} & \dots & \dots & \frac{\exp(ik_f r_{3N}^2 - i\omega(t - \Delta t_{3N}^2))}{r_{3N}^2} \\ \frac{\exp(ik_f r_1^3 - i\omega(t - \Delta t_1^3))}{r_1^3} & \frac{\exp(ik_f r_2^3 - i\omega(t - \Delta t_2^3))}{r_2^3} & \dots & \dots & \frac{\exp(ik_f r_{3N}^3 - i\omega(t - \Delta t_{3N}^3))}{r_{3N}^3} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\exp(ik_f r_1^{N_T} - i\omega(t - \Delta t_1^{N_T}))}{r_1^{N_T}} & \frac{\exp(ik_f r_2^{N_T} - i\omega(t - \Delta t_2^{N_T}))}{r_2^{N_T}} & \dots & \dots & \frac{\exp(ik_f r_{3N}^{N_T} - i\omega(t - \Delta t_{3N}^{N_T}))}{r_{3N}^{N_T}} \end{bmatrix}_{M \times 3N}. \quad (17)$$

Replacing \mathbf{A}_S by $[\mathbf{M}_{SS}]^{-1} \mathbf{V}_S$ in Eq. (16) we get

$$\begin{aligned} \mathbf{P}_T &= \mathbf{Q}_{TS} [\mathbf{M}_{SS}]^{-1} \mathbf{V}_S, \\ \mathbf{V}_T &= \mathbf{M}_{TS} [\mathbf{M}_{SS}]^{-1} \mathbf{V}_S. \end{aligned} \quad (18)$$

D. Pressure field in a homogeneous fluid

Following the DPSM technique described above, the ultrasonic field in a homogeneous fluid (water) generated by a flat circular transducer is computed. For such simple problem geometry the expression of the near field zone length and the divergence angle of the emitted beam can be calculated in closed form (Kundu, 2000). Numerically computed values of

these two parameters are compared with the closed form analytical values to check the accuracy of the numerical results.

From Kundu 2000,

$$N_F = \frac{D^2 - \lambda^2}{4\lambda} \approx \frac{D^2}{4\lambda} \quad \text{for } \lambda \ll D \quad (19a)$$

and

$$\phi = 2 \sin^{-1} \left(1.2 \frac{\lambda}{D} \right), \quad (19b)$$

where N_F is the near field length the distance between the transducer face and the last peak of the pressure field along

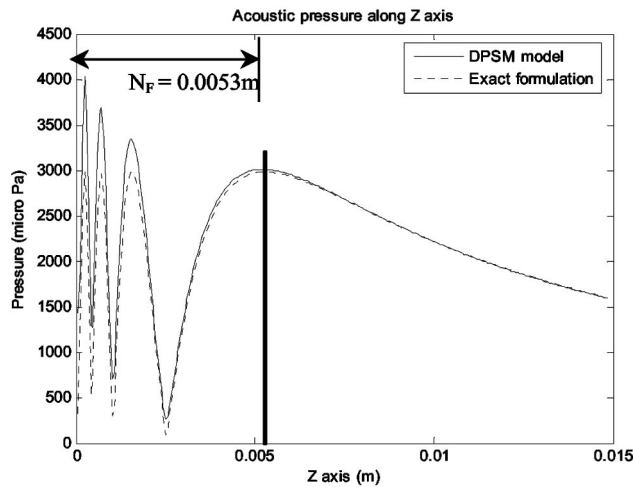


FIG. 3. Acoustic pressure variation along the central axis (Z -axis) of a circular transducer. Coordinate Z is measured from the transducer face. The transducer diameter (D)=2.528 mm, signal frequency (f)=5 MHz. Solid and dashed lines are generated by the DPSM technique and the analytical formula [Eq. (20)], respectively. Near field zone (N_F) predicted by Eq. (19a) matches with the last peak of the curve.

the x_3 (or Z) axis. D is the transducer diameter; λ is the wavelength of the ultrasonic signal, and ϕ is the divergence angle.

The analytical expression of the pressure in a homogeneous fluid along the central axis of a circular transducer is (Placko and Kundu, 2003)

$$p(x_3) = \rho c_f v_0 [\exp(ik_f x_3) - \exp(ik_f \sqrt{x_3^2 + a^2})] \quad (20)$$

Figure 3 shows acoustic pressure variations along the central axis (Z axis) of a circular transducer computed by the DPSM technique described above (solid curve) and the analytical expression [Eq. (20), dashed curve]. Coordinate Z is measured from the transducer face. The transducer area is 2.24 mm², its diameter (D)=2.528 mm, and signal frequency (f)=5 MHz. For the DPSM modeling 128 sources

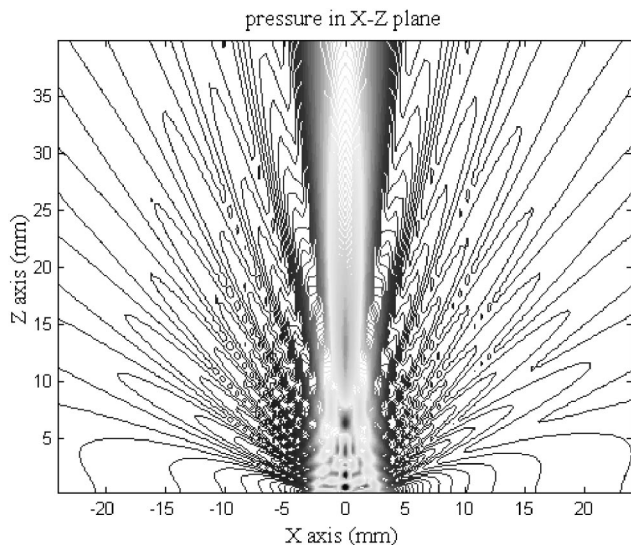


FIG. 4. Ultrasonic field generated by a 6-mm-diam circular transducer excited at 2.2 MHz frequency, following DPSM technique (from Nasser, 2004).

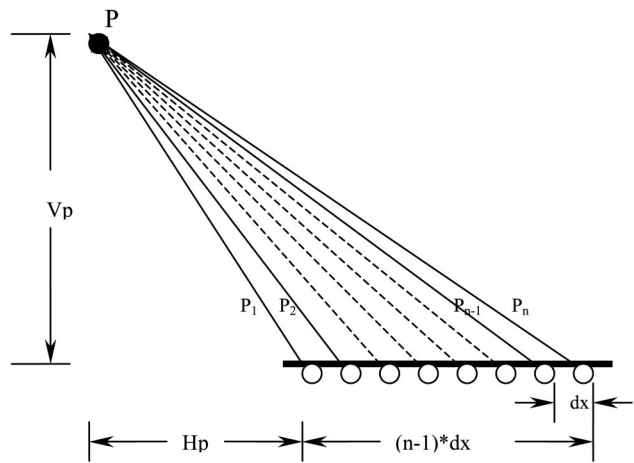


FIG. 5. Focal point P for a phased array transducer. V_p and H_p denote the position of point P relative to the transducer surface.

are placed at $x_3 = -r_s$, while the transducer face is at $x_3 = 0$. For this problem geometry

$$r_s = \sqrt{\frac{\pi D^2}{4 \times 2 \pi N}} = \frac{D}{\sqrt{8N}} = \frac{2.528}{\sqrt{8 \times 128}} = 0.08 \text{ mm.}$$

Since the wave speed in the surrounding fluid (water) is 1.49 km/s, wavelength λ is equal to 0.298 mm for the 5-MHz signal. From Eq. (19a) N_F is found to be 5.3 mm. This theoretical value of N_F matched very well with the last peak. The matching between the DPSM computed curve and the analytical solution in Fig. 3 is also excellent.

Figure 4 shows acoustic pressure in front of a transducer with diameter $D=6$ mm and $f=2.2$ MHz, obtained by the DPSM technique with $N=200$. The theoretical divergence angle ϕ for this transducer is 15.6°, calculated from Eq. (19b). The divergence angle in Fig. 4 is measured and found to be about 15°, very close to the theoretical prediction. These two checks—Figs. 3 and 4—validate the DPSM computed results.

The only restriction of this technique for guaranteeing accurate computation is that the distance between the neighboring point sources should be less than the wavelength of the ultrasonic signal (Placko and Kundu, 2003).

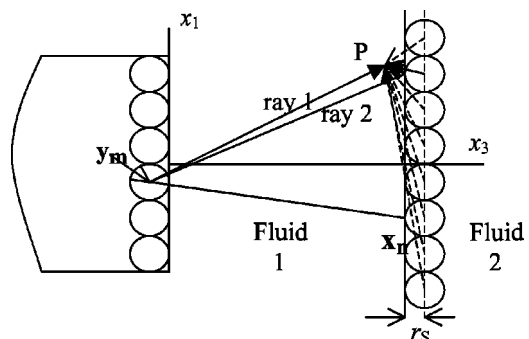


FIG. 6. Point P can receive two rays, direct ray 1 and reflected ray 2, from a single point source y_m .

E. Dynamic focusing and time lag determination

The main purpose of the phased array transducer is to be able to steer the beam and focus it at any point of interest by accurately controlling the time delay of excitation of the point sources. Figure 5 shows how the time delay can be determined for a desired direction of the transmitted beam.

The time lag Δt_n can be obtained from the following equation, so that all signals arrive at point P at the same time:

$$\Delta t_n = \frac{\Delta p}{c} = \frac{P_n - P_{n-1}}{c}, \quad (21)$$

where c is the velocity of the pressure wave in the medium, and P_n and P_{n-1} are the distances of the n th and $(n-1)$ th point sources, respectively, from point P . If the distance between two consecutive point sources is dx , then P_n and P_{n-1} can be expressed as

$$\begin{aligned} P_n &= \sqrt{V_p^2 + [H_p + (n-1)dx]^2} \quad \text{and} \\ P_{n-1} &= \sqrt{V_p^2 + [H_p + (n-2)dx]^2}, \end{aligned} \quad (22)$$

where V_p and H_p are vertical and horizontal distances, respectively, of point P from one edge of the transducer face as shown in Fig. 5.

Note that the selection of the focal point automatically determines the steering direction; thus both the focal point and the steering direction are determined in this manner.

F. Modeling of the ultrasonic field in a homogeneous fluid in the presence of a solid scatterer

When a solid scatterer is placed in front of a transducer in a homogeneous fluid, the ultrasonic field can be computed using the DPSM, by placing point sources both along the transducer surface and along the fluid–solid interface of the solid scatterer. The acoustic field in the fluid is computed by superimposing the contributions of two layers of point sources distributed over the transducer face and the scatterer face as shown in Fig. 6. Two layers of sources are located at a small distance, r_s , away from the transducer face and interface, respectively, such that the apex of every sphere (of radius r_s) touches the transducer face or the fluid–solid interface, as shown in the figure.

Strength of the point sources distributed along the transducer surface can be obtained from Eq. (15) and from the specified boundary conditions at the fluid–solid interface.

For finding the strength of the point sources attached to the interface, velocity components at the interface due to the reflected waves at the interface are to be matched, as described below.

Velocity components at M interface points due to ray 1 only (ignoring reflection) can be easily obtained from the N triplet sources in the following manner [Eq. (16)]:

$$\mathbf{V}_T^i = \mathbf{M}_{TS}^i \mathbf{A}_S, \quad (23)$$

where \mathbf{V}_T^i is the velocity vector at the interface points (\mathbf{x}_n) due to the incident beam only, \mathbf{A}_S is the vector of the point source strengths on the transducer surface. \mathbf{M}_{TS}^i is the matrix that relates the two vectors \mathbf{V}_T^i and \mathbf{A}_S , and the components of \mathbf{M}_{TS}^i are similar to the components of \mathbf{M}_{SS} , the only difference is in the dimensions of the two matrices when the number of point sources at the transducer surface and the interface are different.

For computing the source strength at the interface, the velocity due to the reflected field at the interface (generated by the point sources at the transducer face) and the velocity due to the direct incident field produced by the point sources at the interface are equated. Note that

$$\mathbf{V}_T^r = \mathbf{M}_{TS}^r \mathbf{A}_S \quad (24)$$

is the velocity due to the reflected beam only (ray 2 of Fig. 6). \mathbf{M}_{TS}^r is the $(M \times N)$ [or $(3M \times 3N)$ if all three velocity components are matched] matrix that relates the two vectors \mathbf{V}_T^r and \mathbf{A}_S of Eq. (24). Note that components of \mathbf{M}_{TS}^r can be obtained by multiplying \mathbf{M}_{TS}^i by the appropriate reflection coefficient (Placko and Kundu, 2003).

The velocity field at the same point generated by the interface point sources is obtained from the following equation,

$$\mathbf{V}_T^r = \mathbf{M}_{TI}^i \mathbf{A}_I, \quad (25)$$

where \mathbf{V}_T^r is the $(M \times 1)$ or $[(3M \times 1)]$ vector, same as that in Eq. (24). \mathbf{A}_I is the $(M \times 1)$ or $[(3M \times 1)]$ vector of the strength of interface sources; this vector is unknown. \mathbf{M}_{TI}^i is the $(M \times M)$ or $[(3M \times 3M)]$ matrix that relates the two vectors \mathbf{V}_T^r and \mathbf{A}_I . Components of \mathbf{M}_{TI}^i are computed using similar mathematical formulation as those for \mathbf{M}_{SS} in Eq. (13). However, there are some differences between \mathbf{M}_{SS} and \mathbf{M}_{TI}^i —their dimensions are different and the time delay factor is absent in \mathbf{M}_{TI}^i ,

$$\mathbf{M}_{TI}^i = \begin{bmatrix} f(x_{11}^1, r_1^1) & f(x_{12}^1, r_2^1) & f(x_{13}^1, r_3^1) & \cdots & \cdots & f(x_{1(L-1)}^1, r_{L-1}^1) & f(x_{1L}^1, r_L^1) \\ f(x_{21}^1, r_1^1) & f(x_{22}^1, r_2^1) & f(x_{23}^1, r_3^1) & \cdots & \cdots & f(x_{2(L-1)}^1, r_{L-1}^1) & f(x_{2L}^1, r_L^1) \\ f(x_{31}^1, r_1^1) & f(x_{32}^1, r_2^1) & f(x_{33}^1, r_3^1) & \cdots & \cdots & f(x_{3(L-1)}^1, r_{L-1}^1) & f(x_{3L}^1, r_L^1) \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ f(x_{31}^M, r_1^M) & f(x_{32}^M, r_2^M) & f(x_{33}^M, r_3^M) & \cdots & \cdots & f(x_{3(L-1)}^M, r_{L-1}^M) & f(x_{3L}^M, r_L^M) \end{bmatrix}_{L \times L}, \quad (26)$$

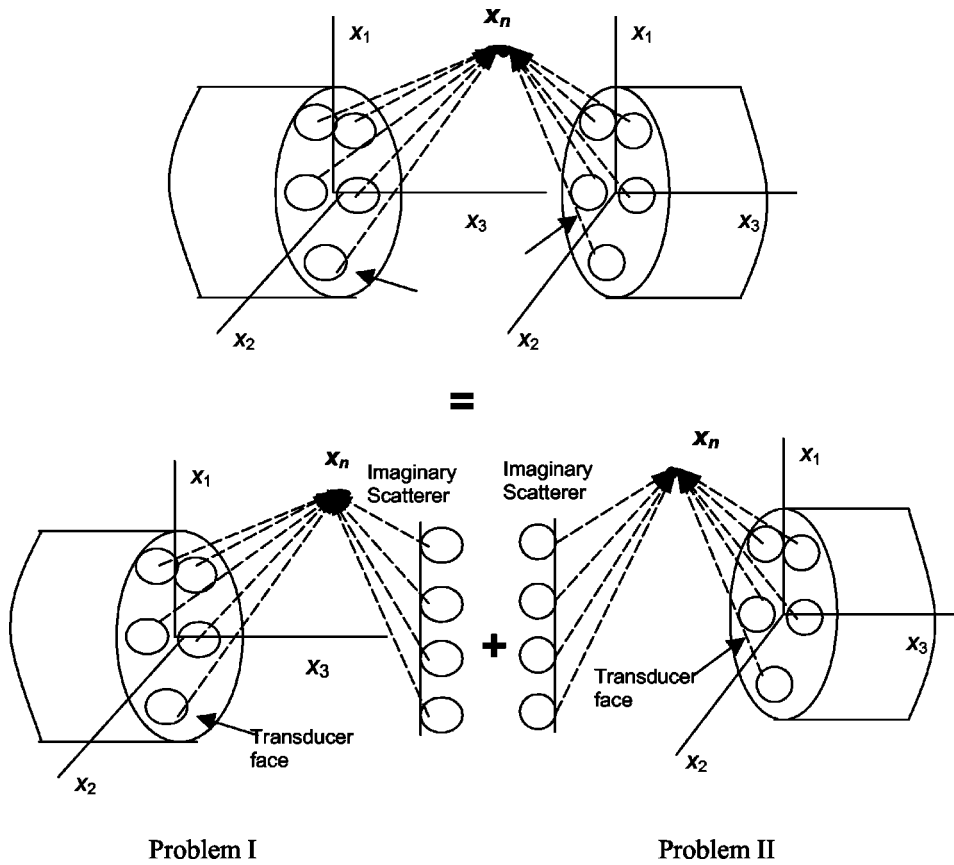


FIG. 7. Modeling of multiple transducers placed in a homogeneous fluid.

where $L=M$ (or $3M$) for matching only one component (or three components) of velocity across the interface,

$$f(x_{jm}^n, r_m^n) = \frac{x_{jm}^n \exp(ik_f r_m^n)}{i\omega\rho(r_m^n)^2} \left(ik_f - \frac{1}{r_m^n} \right). \quad (27)$$

Since \mathbf{A}_I of Eq. (25) is unknown, \mathbf{A}_I is solved in the following manner:

$$\mathbf{A}_I = [\mathbf{M}_{TI}^i]^{-1} \mathbf{V}_T^r = ([\mathbf{M}_{TI}^i]^{-1} \mathbf{M}_{TS}^r) \mathbf{A}_S. \quad (28)$$

The above equation gives the source strength vector at the interface in terms of the source strengths at the transducer face.

The velocity vector on the transducer surface, generated only by the point sources distributed over the transducer face, can be obtained from Eq. (23):

$$\mathbf{V}_S^i = \mathbf{M}_{SS}^i \mathbf{A}_S. \quad (29)$$

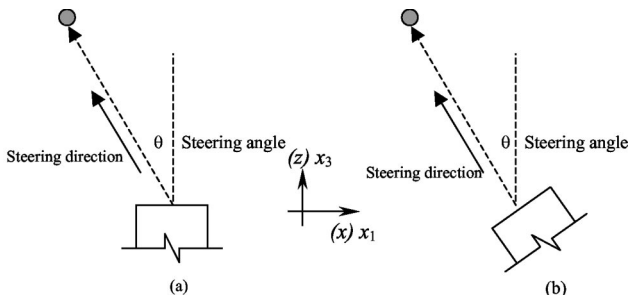


FIG. 8. (a) A phased array transducer with a steering angle θ and (b) a regular transducer rotated at an angle θ .

\mathbf{V}_S^i is the $(N \times 1)$ [or $(3N \times 1)$] vector of the velocity components at the transducer surface. At this step, \mathbf{A}_S is the unknown $(N \times 1)$ [or $(3N \times 1)$] vector of the point source strengths distributed over the transducer face and \mathbf{M}_{SS}^i is the $(N \times N)$ [or $(3N \times 3N)$] matrix, identical to the matrix \mathbf{M}_{SS} , defined in Eq. (13).

In the same manner, velocity components on the transducer surface, only due to the interface sources are obtained,

$$\mathbf{V}_S^r = \mathbf{M}_{SI}^i \mathbf{A}_I. \quad (30)$$

The above equation is derived from Eq. (25), after placing the target points on the transducer surface. Here, \mathbf{V}_S^r is a $(N \times 1)$ [or $(3N \times 1)$] vector of the velocity components at N points on the transducer surface, \mathbf{A}_I is the unknown $(M \times 1)$ [or $(3M \times 1)$] vector of the interface source strengths and \mathbf{M}_{SI}^i is the $(N \times M)$ [or $(3N \times 3M)$] matrix, similar to the one given in Eq. (13).

Adding Eqs. (29) and (30), the total velocity at the transducer surface is obtained:

$$\mathbf{V}_S = \mathbf{V}_S^i + \mathbf{V}_S^r = \mathbf{M}_{SS}^i \mathbf{A}_S + \mathbf{M}_{SI}^i \mathbf{A}_I. \quad (31)$$

Substituting Eqs. (24) and (25) into Eq. (31):

$$\begin{aligned} \mathbf{V}_S &= \mathbf{M}_{SS}^i \mathbf{A}_S + \mathbf{M}_{SI}^i \mathbf{A}_I \\ &= \mathbf{M}_{SS}^i \mathbf{A}_S + \mathbf{M}_{SI}^i [\mathbf{M}_{TI}^i]^{-1} \mathbf{M}_{TS}^r \mathbf{A}_S \Rightarrow \mathbf{V}_S \\ &= [\mathbf{M}_{SS}^i + \mathbf{M}_{SI}^i [\mathbf{M}_{TI}^i]^{-1} \mathbf{M}_{TS}^r] \mathbf{A}_S, \end{aligned} \quad (32)$$

or,

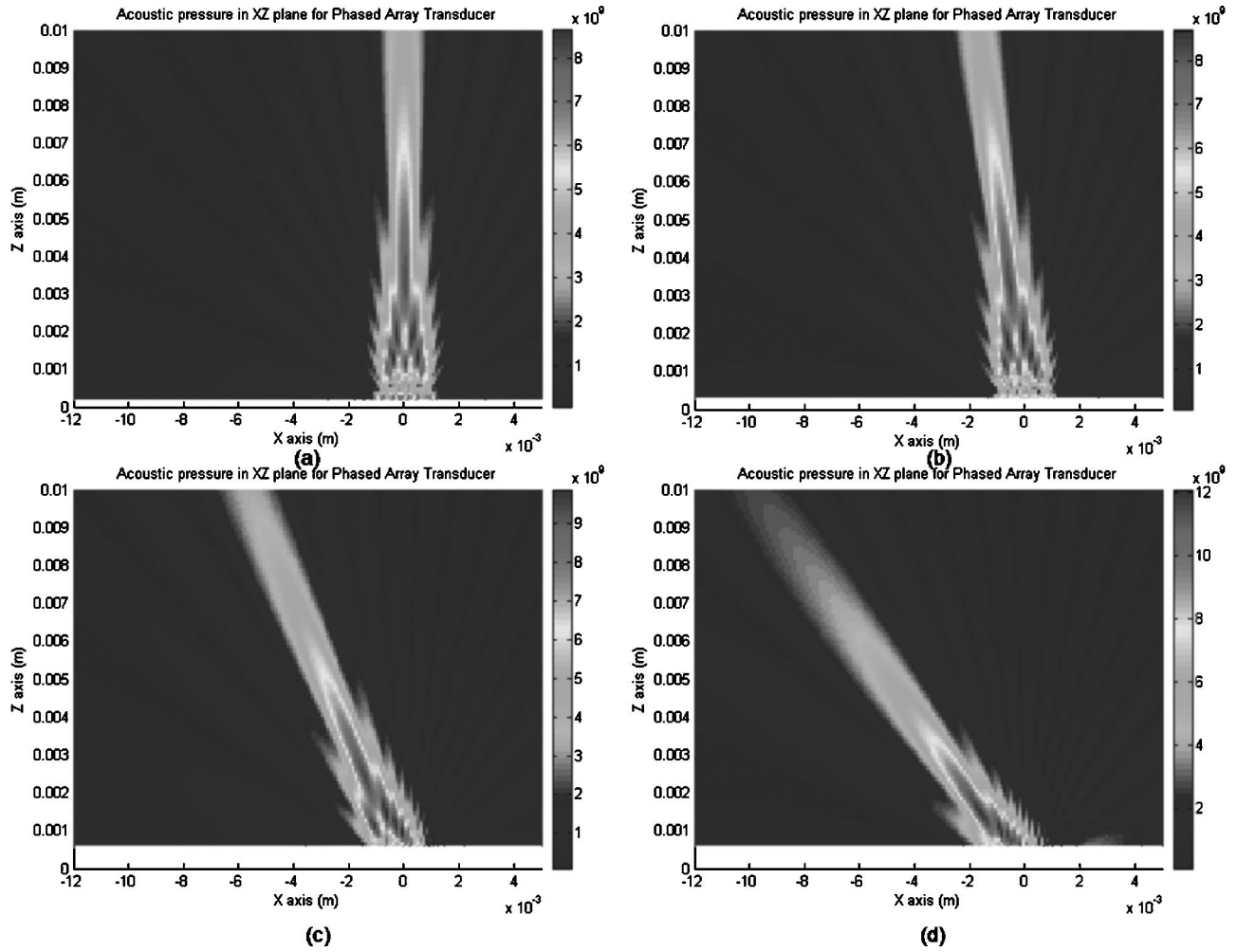


FIG. 9. Ultrasonic pressure fields for a phased array transducer for steering angles equal to (a) 0°, (b) 10°, (c) 30° and (d) 45°.

$$\mathbf{A}_S = [\mathbf{M}_{SS}^i + \mathbf{M}_{SI}^i [\mathbf{M}_{TI}^i]^{-1} \mathbf{M}_{TS}^r]^{-1} \mathbf{V}_S, \quad (33)$$

where

$$\mathbf{V}_S = [0 \ 0 \ v_0 \ 0 \ 0 \ v_0 \ 0 \ 0 \ v_0]^T. \quad (34)$$

Note that v_0 is the value of v_3 (x_3 component of velocity) on the transducer surface. This choice for \mathbf{V}_S reflects the fact that the vibration produced by a piezoelectric transducer has only normal velocity components on its active surface.

After obtaining \mathbf{A}_S , the pressure field at any point between the transducer and the scatterer can be computed in the following manner:

$$\begin{aligned} \mathbf{P}_T &= \mathbf{P}_S^i + \mathbf{P}_I^i, \\ \mathbf{P}_S^i &= \mathbf{Q}_{TS}^i \mathbf{A}_S = \mathbf{Q}_{TS}^i [\mathbf{M}_{SS}^i + \mathbf{M}_{SI}^i [\mathbf{M}_{TI}^i]^{-1} \mathbf{M}_{TS}^r]^{-1} \mathbf{V}_S, \end{aligned} \quad (35)$$

TABLE I. Water and steel properties.

Material and properties	P-wave speed (km/s)	S-wave speed (km/s)	Density (gm/cc)
Steel	5.96	3.26	7.932
Water	1.49	...	1

$$\begin{aligned} \mathbf{P}_I^i &= \mathbf{Q}_{TI}^i \mathbf{A}_I = \mathbf{Q}_{TI}^i [[\mathbf{M}_{TI}^i]^{-1} \mathbf{M}_{TS}^r] \mathbf{A}_S \\ &= \mathbf{Q}_{TI}^i [[\mathbf{M}_{TI}^i]^{-1} \mathbf{M}_{TS}^r] [\mathbf{M}_{SS}^i + \mathbf{M}_{SI}^i [\mathbf{M}_{TI}^i]^{-1} \mathbf{M}_{TS}^r]^{-1} \mathbf{V}_S. \end{aligned}$$

G. Interaction between two transducers in a homogeneous fluid

When two transducers are placed in a homogeneous fluid, the ultrasonic field can be modeled by superimposing two simpler problems as shown in Fig. 7. Fields due to these two simpler problems can be added to obtain the total ultrasonic field at any point x_n . Note that each transducer face will act as the source of energy as well as a reflecting surface. For simplicity we will call one of the transducers as the left side transducer and the second one as the right side transducer (see Fig. 7).

As shown in Fig. 7, the original problem is decomposed into two problems—in one problem, the right transducer is replaced by an imaginary scatterer and, in the second problem, the left transducer is replaced by the imaginary scattering surface (see the bottom two figures of Fig. 7). For the first problem (the left figure), the ultrasonic field at a point \mathbf{x}_n is the combined effect of the ultrasonic rays coming from the left side transducer face and the imaginary scatterer on the right side. For the second problem (the right figure) the

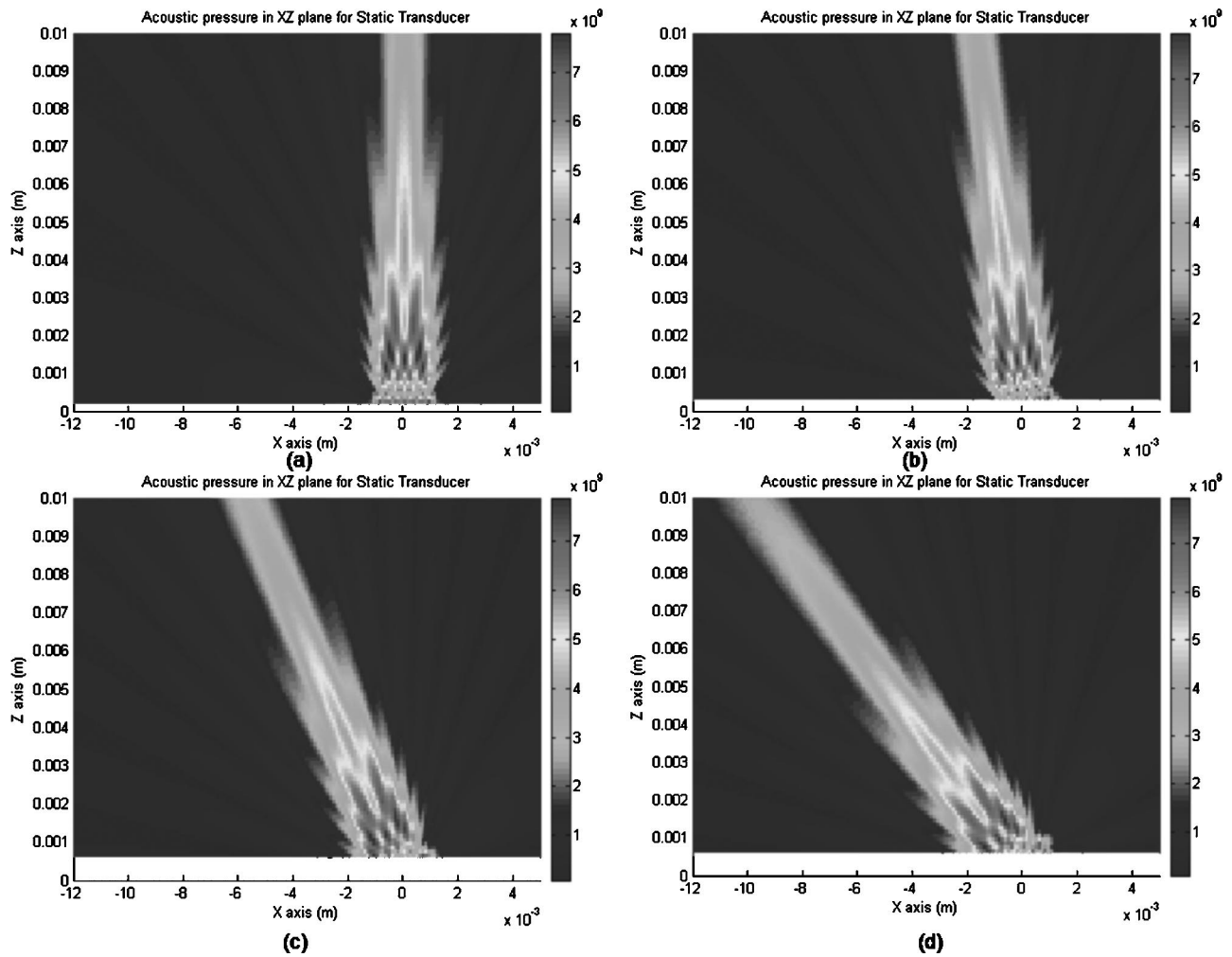


FIG. 10. Ultrasonic pressure fields for a conventional planar transducer for rotation angles equal to (a) 0° , (b) 10° , (c) 30° and (d) 45° .

acoustic field at point \mathbf{x}_n is the combined effect of the right side transducer and the imaginary scatterer on the left side.

For both problems the transducer and the scatterer surfaces are modeled as distributed point sources. The total ultrasonic field at point \mathbf{x}_n generated by the two transducers is then obtained by superimposing the two solutions. In mathematical notations, the modeling steps described above can be written as

$$\begin{aligned}
 \mathbf{P}_{left} &= \mathbf{P}_{S(left)}^i + \mathbf{P}_{I(right\ scatterer)}^i, \\
 \mathbf{P}_{right} &= \mathbf{P}_{S(right)}^i + \mathbf{P}_{I(left\ scatterer)}^i, \\
 \mathbf{P}_{Total} &= \mathbf{P}_{left} + \mathbf{P}_{right},
 \end{aligned} \tag{36}$$

where \mathbf{P}_{left} and \mathbf{P}_{right} are the pressure fields generated by problems 1 and 2, respectively. \mathbf{P}_S^i and \mathbf{P}_I^i are obtained from Eq. (35). \mathbf{P}_{Total} is the combined effect for the two transducers. Note that \mathbf{P}_I^i takes into account the interaction effect. If the ultrasonic fields generated by the two transducers are simply added without taking into account the interaction (or scattering) effect of one transducer on the other, then \mathbf{P}_I^i will not appear in the formulation.

III. NUMERICAL RESULTS AND DISCUSSION

The main objective of this paper is to model a phased array transducer with dynamic focusing using the distributed point source method (DPSM). The first part of this research work involves the modeling of the time-dependent excitation of point sources in a phased array transducer. The second part focuses on studying the interaction effect when two phased array transducers are placed, face to face, in a homogeneous fluid (water) medium. Phased array transducers provide dynamic capabilities compared to the conventional static transducers. For conventional transducers the transducer surfaces are rotated to achieve the desired steering angle (see Fig. 8). In a phased array transducer, the transducer surface does not have to be rotated or translated (moved). For a phased array transducer, different segments of the transducer are activated at different time intervals. In the DPSM modeling technique individual point sources are excited at a predetermined time interval. MATLAB codes, based on the DPSM, have been developed to introduce the time-dependent excitation of the point sources.

A. Dynamic steering and focusing

As mentioned earlier, in a phased array transducer, dynamic steering and focusing can be achieved by suitably se-

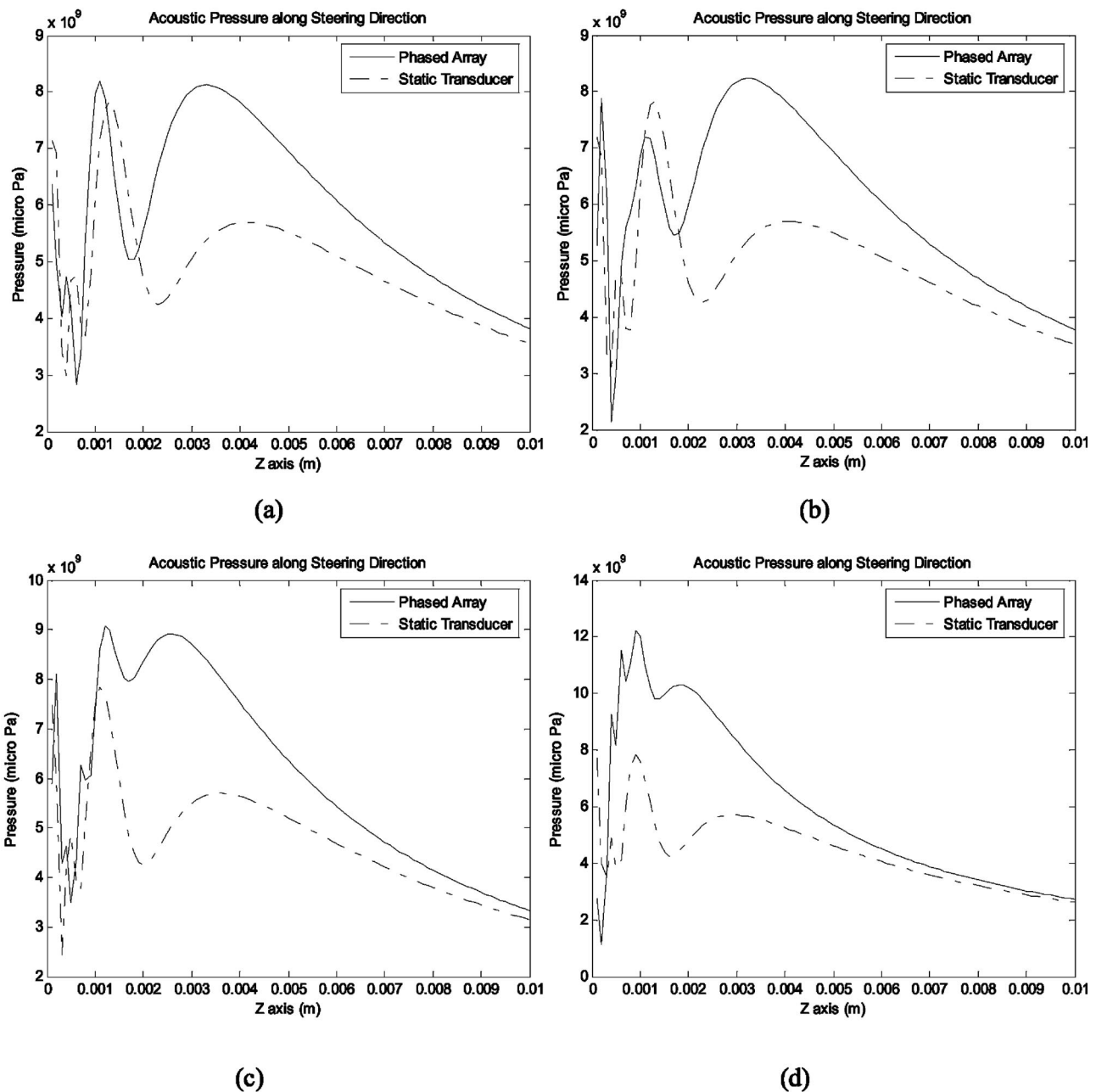


FIG. 11. Comparison of acoustic pressures along the steering directions generated by the phased array transducer and the conventional static transducer for steering angles equal to (a) 0° , (b) 10° , (c) 30° and (d) 45° . The pressure values are proportional to the transducer surface velocity amplitude v_0 and can be easily changed by simply altering v_0 . For this reason the absolute values of the plotted pressure are not very important; however, the variation of the pressure, which is independent of the transducer surface velocity amplitude, is important.

lecting the time phase and point of focus as described in Sec. II E. In the first investigation, we modeled a phased array transducer that produces ultrasonic beams at angles $\sim 0^\circ$, 10° , 30° and 45° . Figure 9 shows the pressure fields generated by a rectangular phased array transducer in a homogeneous fluid (water). The area of the transducer surface is 2.5 mm^2 and its length to width ratio is 2:1. The outermost layer of the transducer is taken as steel. The acoustic properties of steel and water are given in Table I. We will use the term “steering angle” for the angle that the ultrasonic beam makes with the normal direction to the transducer face. The excitation frequency is 5 MHz.

The purpose of this investigation is to compare the ultrasonic fields generated by phased array transducers and

conventional planar transducers rotated at different steering angles. Figures 9(a)–(d), show the acoustic pressure fields generated by a phased array transducer for steering angles equal to $\sim 0^\circ$, 10° , 30° and 45° , respectively. Note that the scales in x and y directions are different. Figures 10(a)–(d), show the acoustic pressures generated by a conventional planar transducer when the axis of the transducer surface is rotated by 0° , 10° , 30° and 45° angles, respectively. In this paper phased array and conventional transducers are also called as dynamic and static transducers, respectively. For both these transducers, the surface area is equal to 2.5 mm^2 and the number of point sources is equal to 98. Fourteen rows of elements are distributed along the x direction representing 14 physical elements of the phased array transducer.

Each physical element is discretized into 7 point sources distributed along the y direction, giving rise to a total of 98 point sources. Note that the spacing between two neighboring point sources is smaller than the wavelength. Such small spacing is needed for right convergence of the solution (Placko and Kundu, 2003).

In Fig. 9, the normal distance of the focal point from the transducer face is set at 10 mm in the Z direction. The steering angles of 0° , 10° , 30° and 45° are achieved by changing the horizontal distance of the focus point, keeping the normal distance fixed at 10 mm. Figures 9(a) and 10(a) are pressure fields generated by the phased array and conventional static transducers for a steering angle of 0° . From these figures, it can be seen that the maximum pressure intensities are 8.2×10^9 and 7.6×10^9 units for phased array and static transducers, respectively. In the figures, it is also observed that the pressure beam produced by the phased array transducer is narrower or more collimated than that produced by the conventional static transducer. Similar conclusions are drawn from Figs. 9(b) and 10(b) for 10° steering angle, Figs. 9(c) and 10(c) for 30° steering angle and Figs. 9(d) with 10(d) for 45° steering angle. Such observation is expected because the phased array transducer is focused at a finite distance while the conventional planar transducer is focused at infinity.

Figures 11(a)–(d), plot the pressure fields along the steering directions for both phased array and conventional transducers for 0° , 10° , 30° and 45° inclination angles, respectively. These figures show that near the transducer face, the pressure intensity for the two cases are somewhat close, but beyond this near field region the pressure difference between these two types of transducers is much greater. The phased array transducer produces stronger pressure beams than the static transducer in the region between the focal point ($Z = 10$ mm) and the near field distance N_F .

B. Interaction between two phased array transducers placed face to face

In the second investigation, the interaction effect between two phased array transducers placed in a homogeneous medium is studied. The transducer faces are placed parallel to each other as shown in Fig. 12. The central axes of the transducers are collinear. The surrounding medium is water. The direct field generated by the first transducer as well as the scattered field from the second transducer is calculated. Note that each transducer surface is modeled as the energy-generating surface as well as the scattering surface or the reflecting surface for the ultrasonic field generated by the other transducer.

The steering angle for the phased array transducer is varied to study the interaction effect as a function of the steering angle. We studied the interaction of the transducers for three different steering angles. For the first case, the steering angle is very small ($\theta \approx 0^\circ$). For the second and third cases the steering angle is increased to 10° and 30° , respectively. The reason for the incident angle to be close to 0° but not exactly equal to 0° is that during the discretization process of the transducer face the point source placements have some randomness. Thus for the 0° case all point sources are not necessarily placed in perfect symmetry with

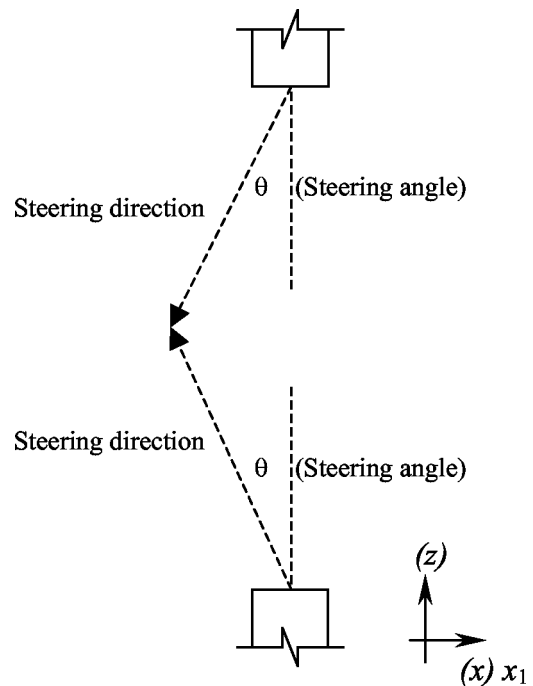


FIG. 12. Steering direction for two transducers.

respect to the transducer axis going through the focal point. Therefore, the field generated for the 0° case is not necessarily perfectly symmetric about the central axis of the transducer. Three different steering angles (θ) are analyzed because when θ is small (close to 0°), the bulk of the incident pressure beam hits the surface of the second transducer face. For 10° steering angle, a part of the beam hits the opposite transducer and, for 30° steering angle, the bulk of the incident beam travels through the fluid medium without hitting the second transducer. We compared the acoustic pressure fields generated by the phased array transducers when the interaction or the scattering effect of the second transducer is ignored and when it is included. Figures 13(a) and (b) show the pressure fields generated by the phased array transducers for a small steering angle ($\theta \approx 0^\circ$) when the interaction effect is ignored. Figures 14(a) and (b) plot the acoustic pressures generated by the same transducers when the interaction effect is included. Let us state here again what we mean by the interaction effect being ignored and included. The acoustic pressure field in Fig. 13(a) is generated by the bottom transducer when the top transducer is absent. In other words, in Fig. 13(a) the ultrasonic waves generated by the bottom transducer propagate freely without being interrupted or scattered by the top transducer. The top transducer generates a similar beam that is simply a mirror image of the bottom transducer generated beam and is not shown here. Figure 13(b) is the superposition of uninterrupted fields generated by the bottom transducer [Fig. 13(a)] and the top transducer (not shown).

On the other hand, Figs. 14(a) and (b) show the pressure fields obtained after considering the interaction effect. Figure 14(a) shows the pressure field generated by the bottom transducer when the top transducer, placed at a distance of 10 mm, acts as a scatterer, in this case mostly as a reflector. The pressure beam generated by the bottom transducer is partly

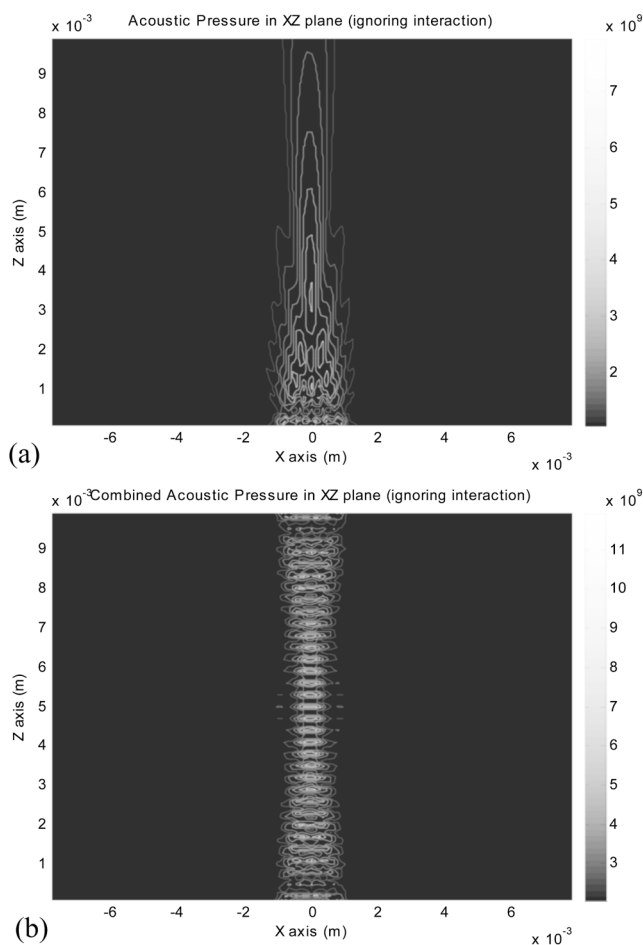


FIG. 13. Acoustic pressure in the XZ (or x_1x_3) plane for phased array transducers when the interaction effect is ignored. Transducers are placed at a distance of 10 mm. (a) shows pressure fields generated by the bottom transducer in the absence of the top transducer when the steering angle is 0° . (b) shows the total field, obtained by adding two fields generated by the bottom transducer [shown in (a)] and the top transducer [a mirror image of (a)].

reflected or scattered by the top transducer and thus affects the total pressure field in the homogeneous medium. It should be mentioned here that for Fig. 14(a) the top transducer surface acts only as a scatterer and not as a wave generator. The top transducer generates a similar beam that is simply a mirror image of the bottom transducer generated beam and is not shown here. Figure 14(b) is the superposition of the field shown in Fig. 14(a) and a similar field (simply a mirror image) generated by the top transducer. The steering angle is $\sim 0^\circ$ for both Figs. 13 and 14.

In Fig. 14(a), the interaction effect (due to the reflection at the top transducer surface) can be clearly seen. In this figure the interference between the direct incident ultrasonic beam from the bottom transducer and the reflected beam from the top transducer causes alternate peaks and dips in the pressure field whereas, in Fig. 13(a), no such peaks and dips are observed. The difference between Figs. 13(b) and 14(b) is comparatively less prominent; however, one can see from these two figures that in the central region, near $z = 5$ mm, the beam in Fig. 14(b) is slightly wider than that in Fig. 13(b).

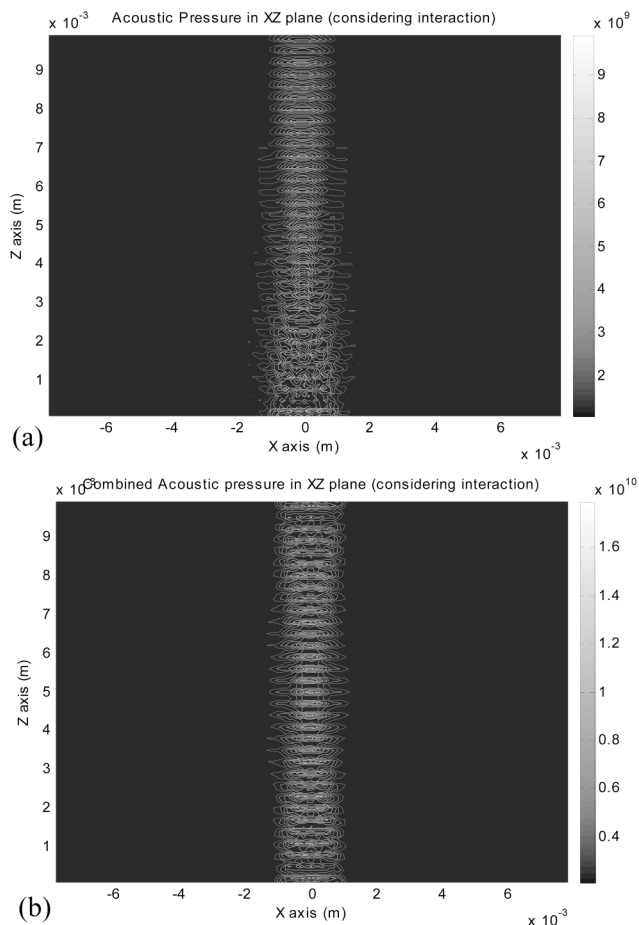


FIG. 14. Acoustic pressure in the XZ (or x_1x_3) plane for phased array transducers when the interaction effect is considered. Transducers are placed at a distance of 10 mm. (a) shows pressure fields generated by the bottom transducer when the top transducer acts as a scatterer or reflector, for 0° steering angle. (b) shows the total field, obtained by adding two fields generated by the bottom transducer [shown in (a)] and the top transducer [a mirror image of (a)].

In Fig. 15(a) the pressure variations along the steering direction are plotted for both cases, i.e., when the interaction effect is considered and when it is ignored. In Fig. 15(a), the two lines almost coincide. When the difference between the two pressure fields is plotted one can see in Fig. 15(b) that the difference is more prominent near the transducer faces. Figure 15(c) shows the pressure difference as a percentage of the total pressure plotted as a function of the distance from the transducer surface. Note that the maximum difference between the two pressure fields is about 4% within 1 mm of the transducer face and it is less than 0.5% when the distance from the transducer face is greater than 2.5 mm. The reason for the pressure difference being higher near the transducer face is that the reflected pressure produced by the second transducer is relatively stronger near the reflecting surface and it becomes weaker as the distance from the reflecting surface increases.

Figure 16 shows the pressure field when the steering angle is 10° and the interaction effect is considered. Figure 16(a) plots the pressure field generated by the bottom transducer and then scattered by the top transducer. Small scattering effect can be noticed in the top part of the beam for z

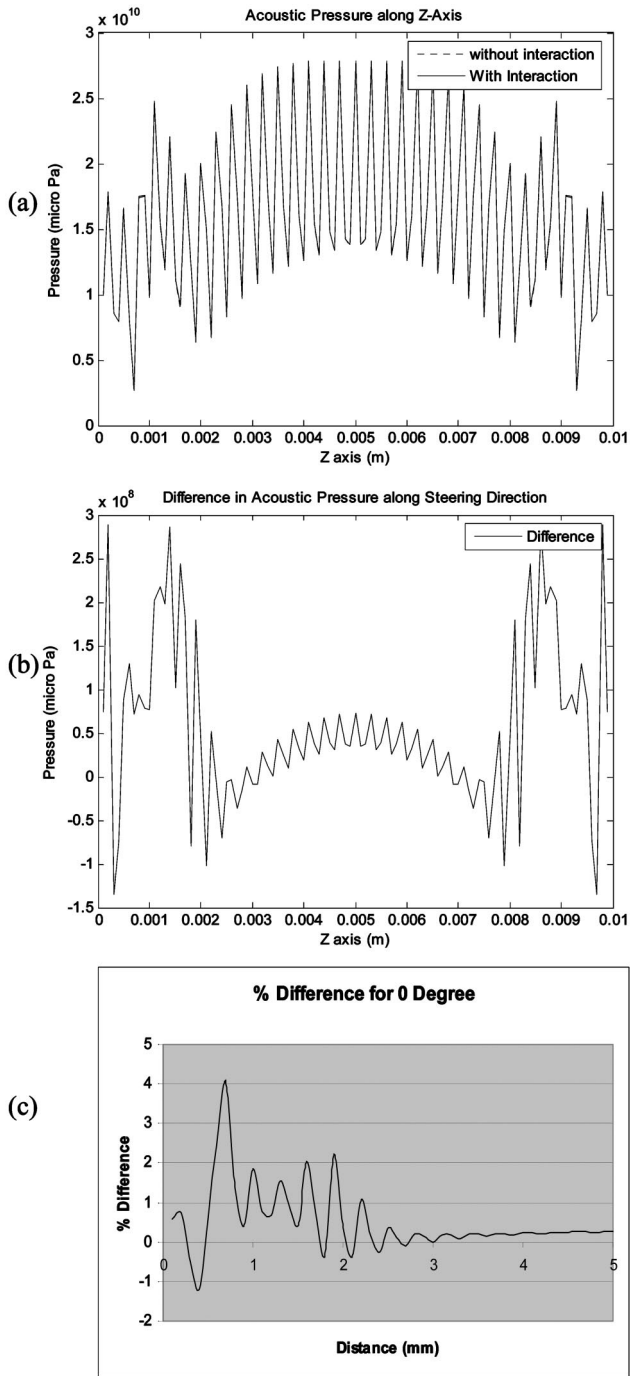


FIG. 15. Acoustic pressures generated by two phased array transducers, placed face to face with steering angle = 0° , when the interaction effect is considered and when it is ignored. (a) shows the two pressure fields that almost coincide along the central axis of the transducers. (b) shows the difference between the two pressure fields. (c) shows the percentage difference between the two fields. Because of symmetry the percentage difference is plotted from the transducer face to the central plane only.

between 9 and 10 mm. Figure 16(b) shows the total effect when both top and bottom transducers are excited. The pressure fields generated after ignoring the interaction effect are very similar to the figures shown in Fig. 16 although not identical, and not shown separately here. The difference between these two pressure fields is plotted in Fig. 17. Figure 17(a) shows the difference between the total pressure fields along the steering line when the interaction effect is consid-

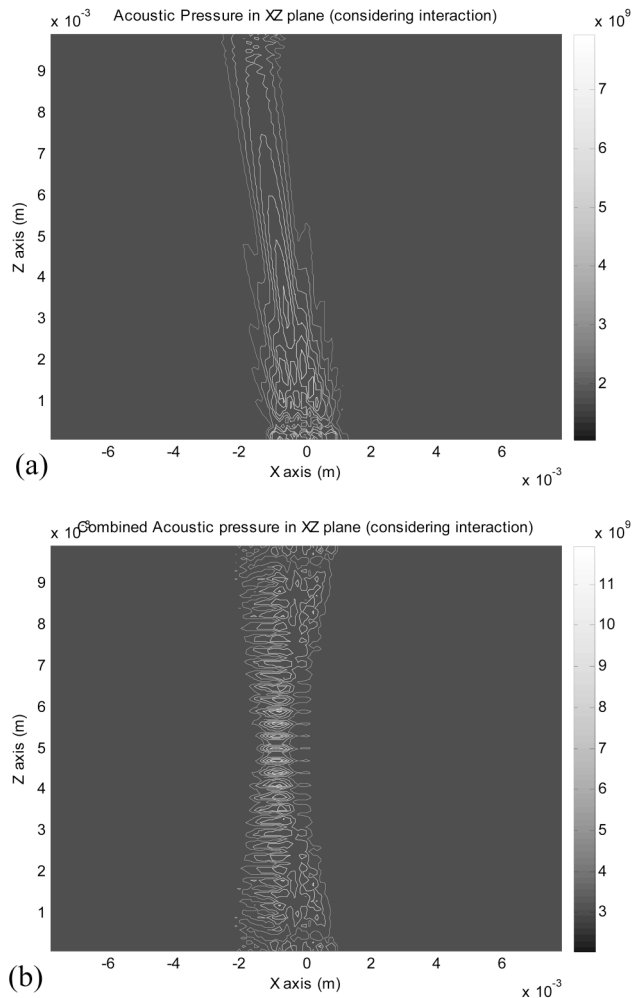
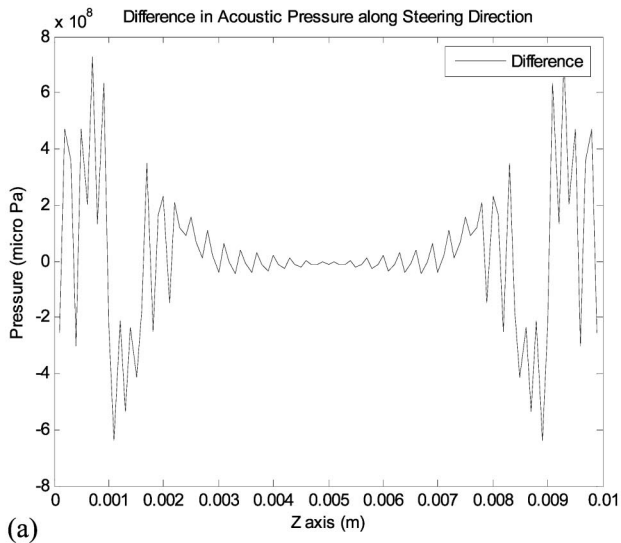
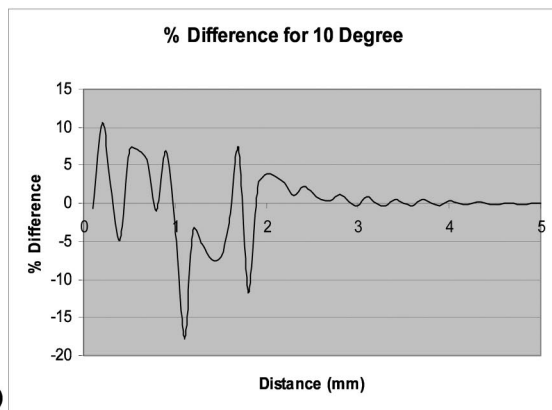


FIG. 16. Acoustic pressure in the XZ (or x_1x_3) plane for phased array transducers when the interaction effect is considered. Transducers are placed at a distance of 10 mm. (a) shows pressure fields generated by the bottom transducer when the top transducer acts as a scatterer or reflector, for 10° steering angle. (b) shows the total field, obtained by adding two fields generated by the bottom transducer [shown in (a)] and the top transducer [a mirror image of (a)].

ered [Fig. 16(b)] and when it is ignored (corresponding figure is not shown); Fig. 17(b) shows the percentage difference between the two pressure fields along the steering line. Because of symmetry the percentage difference is plotted from the transducer face to the central plane only. Note that the maximum difference between the two pressure fields is about 17.5% within 1.4 mm of the transducer face and it is less than 0.2% when the distance from the transducer face is greater than 2.8 mm. From Figs. 16 and 17, it is clear that the interaction from the second transducer surface has some effect on the total pressure distribution. It is interesting to note that although for the inclined beam comparatively less energy is being reflected by the second transducer the difference between the two pressure fields is greater near the transducer face in Fig. 17(b) in comparison to that in Fig. 15(c). For $\sim 0^\circ$ steering angle probably some destructive interference between the two reflected beams reduces the interaction effect to some extent. However, in the central region Fig. 15(c) shows a higher value in comparison to that in Fig. 17(b).



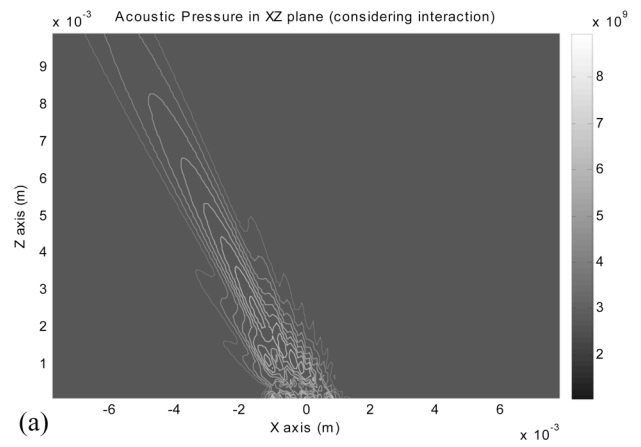
(a)



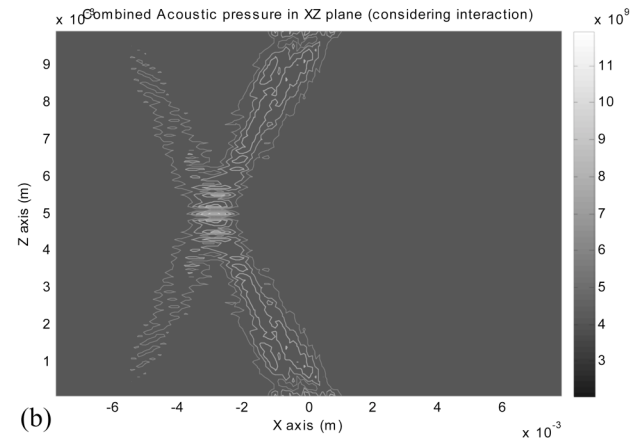
(b)

FIG. 17. Pressure difference for 10° steering angle. (a) Difference between the total pressure fields along the steering line when the interaction effect is considered [Fig. 16(b)] and when it is ignored (not shown); (b) shows the percentage difference between the two pressure fields along the steering line. Because of symmetry the percentage difference is plotted from the transducer face to the central plane only.

Figure 18 shows the pressure field for 30° steering angle when the interaction effect is considered. These plots are similar to Fig. 16, the only difference is that the steering angle for Fig. 18 is 30° and for Fig. 16 it was 10° . Figure 18(a) plots the pressure field generated by the bottom transducer and scattered by the top transducer. Hardly any scattering effect can be observed in this figure because the pressure beam almost completely misses the second transducer. Figure 18(b) shows the total effect when both top and bottom transducers are excited. Pressure fields generated by ignoring the interaction effect looks almost identical to the figures shown in Fig. 18 and not shown separately here. Although to the naked eye they look almost identical, there is some difference between these two pressure fields. This difference is plotted in Fig. 19. Figure 19(a) shows the difference between the total pressure fields along the steering line when the interaction effect is considered [Fig. 18(b)] and when it is ignored (corresponding figure is not shown); Fig. 19(b) shows the percentage difference between the two pressure fields along the steering line. Because of symmetry the percentage difference is plotted from the transducer face to the central



(a)



(b)

FIG. 18. Acoustic pressure in the XZ (or x_1x_3) plane for phased array transducers when the interaction effect is considered. Transducers are placed at a distance of 10 mm. (a) shows pressure fields generated by the bottom transducer when the top transducer acts as a scatterer or reflector, for 30° steering angle. (b) shows the total field, obtained by adding two fields generated by the bottom transducer [shown in (a)] and the top transducer [a mirror image of (a)].

plane only. Note that the maximum difference between the two pressure fields is about 14% within 1 mm of the transducer face and it is less than 0.1% when the distance from the transducer face is greater than 1.2 mm. In the central zone, the pressure difference is negligible. Therefore, even when the ultrasonic beam from one transducer apparently misses the other transducer, there is some significant scattering effect near the transducer face and it should not be ignored while computing the total field.

IV. CONCLUSION

The advantage of using a phased array transducer is its faster scanning capability that is achieved by removing the mechanical movement of the transducer and combining the electronic focusing with the beam steering. In this research, the distributed point source method (DPSM) is used to model phased array transducers. Different point sources placed on the transducer face are excited with varying time delays. The delay in excitation time is changed to produce acoustic beams with various steering angles. Comparison between the pressure fields produced by phased array transducers and conventional transducers for different steering angles shows that the focusing effect in the phased array transducers helps

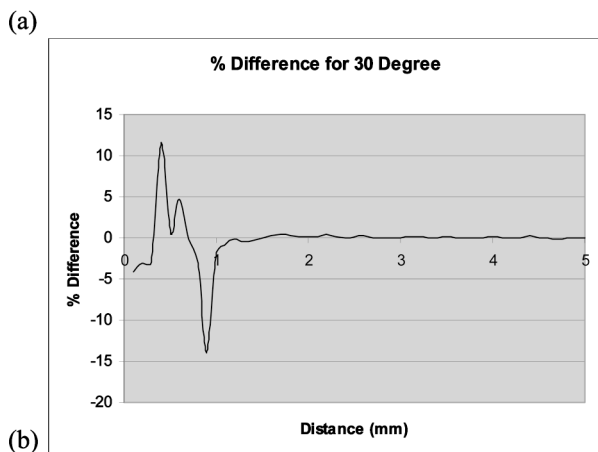
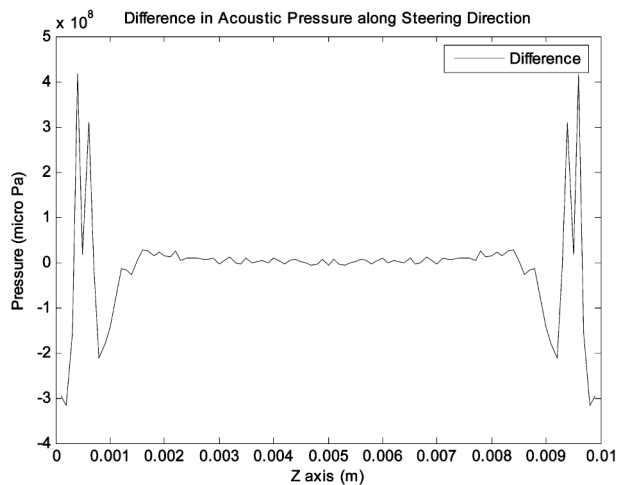


FIG. 19. Pressure difference for 30° steering angle. (a) Difference between the total pressure fields along the steering line when the interaction effect is considered [Fig. 18(b)] and when it is ignored (not shown); (b) shows the percentage difference between the two pressure fields along the steering line. Because of symmetry the percentage difference is plotted from the transducer face to the central plane only.

us in producing stronger and better collimated beams than the conventional planar transducers. Pressure fields are modeled for single transducers as well as for pairs of phased array transducers placed face to face, to study the interaction effect for different steering angles. As expected, relatively stronger interaction effect is observed near the transducer face because the scattered field is relatively stronger there. Interestingly, even when the steering angle is large and the beam generated by one transducer apparently misses the other transducer, the field near the transducer face produced by the simple superposition of two generated beams significantly differs from the total field that includes the interaction effect. However, in the central region, away from the transducer face, the interaction or scattering effect is very small. Therefore, simple superposition of two ultrasonic fields generated by two individual phased array transducers working independently, in absence of the second transducer, is sufficient to model the combined ultrasonic field in this region.

ACKNOWLEDGMENTS

Partial financial supports from the National Science Foundation (Grant Nos. INT-9912549 and CMS-9901221),

USA, and that of ENS Cachan, France, to the first two authors for their visits to France to carry out this research are gratefully acknowledged.

- Ahmad, R., Kundu, T., and Placko, D. (2004). "Modeling of the Ultrasonic Field of Two Transducers Immersed in a Homogeneous Fluid Using Distributed Point Source Method," *I2M (Instrum., Measur. Metrol.) J.* **3**, 87–116.
- Azar, L., and Wooh, S. C. (1999a). "A novel ultrasonic phased arrays for the nondestructive evaluation of concrete structures," in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York), Vol. 18B, pp. 2153–2160.
- Azar, L., and Wooh, S. C. (1999b). "Phase steering and focusing behavior of ultrasound in cementitious materials," in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York), Vol. 18B, pp. 2161–2167.
- Kundu, T. (2000). "Chapter 12: Nondestructive Testing Techniques for Material Characterization," in *Modeling in Geomechanics*, edited by M. Zaman, G. Gioda, and J. Booker (Wiley, New York), pp. 267–298.
- Lee, J. P., Placko, D., Alnuamimi, N., and Kundu, T. (2002). "Distributed Point Source Method (DPSM) for Modeling Ultrasonic Fields in Homogeneous and Non-Homogeneous Fluid Media in Presence of an Interface," *First European Workshop on Structural Health Monitoring*, edited by D. Balageas (DEStech, Lancaster, PA), pp. 414–421, Ecole Normale Supérieure de Cachan, France, 10–12 July.
- Nasser, A. (2004). "Modeling Ultrasonic Transducers in Homogeneous and Non-Homogeneous Media Using DPSM Method," Ph.D. dissertation, Department of Civil Engineering and Engineering Mechanics, University of Arizona, Tucson, AZ.
- Placko, D., and Kundu, T. (2001). "A Theoretical Study of Magnetic and Ultrasonic Sensors: Dependence of Magnetic Potential and Acoustic Pressure on the Sensor Geometry," in *Advanced NDE for Structural and Biological Health Monitoring*, edited by T. Kundu, SPIE's 6th Annual International Symposium on NDE for Health Monitoring and Diagnostics, 4–8 March 2001, Newport Beach, CA, Vol. 4335, pp. 52–62.
- Placko, D., and Kundu, T. (2003). "Chapter 2: Modeling of Ultrasonic Field by Distributed Point Source Method," in *Ultrasonic Nondestructive Evaluation: Engineering and Biological Material Characterization*, edited by T. Kundu (CRC, Boca Raton, FL), pp. 143–202.
- Placko, D., Kundu, T., and Ahmad, R. (2001). "Distributed Point Source Method (DPSM) for Modeling Ultrasonic Fields in Homogeneous and Non-Homogeneous Fluid Media," *Second International Conference on Theoretical, Applied, Computational and Experimental Mechanics*, Kharagpur, India, 27–30 December 2001 (Globenet Computers, India), published in CD, Paper No. 114.
- Placko, D., Kundu, T., and Ahmad, R. (2002). "Theoretical Computation of Acoustic Pressure Generated by Ultrasonic Sensors in Presence of an Interface," in *Smart NDE and Health Monitoring of Structural and Biological Systems*, edited by T. Kundu, SPIE's 7th Annual International Symposium on NDE for Health Monitoring and Diagnostics, 18–21 March 2002, San Diego, CA, Vol. 4702, pp. 157–168.
- Placko, D., Kundu, T., and Ahmad, R. (2003). "Ultrasonic Field Computation in Presence of a Scatterer of Finite Dimension," in *Smart Nondestructive Evaluation and Health Monitoring of Structural and Biological Systems*, edited by T. Kundu, SPIE's 8th Annual International Symposium on NDE for Health Monitoring and Diagnostics, 3–5 March 2003, San Diego, CA, Vol. 5047, pp. 169–179.
- Placko, D., Liebeaux, N., and Kundu, T. (2001). "Presentation d'une methode generique pour La Modelisation des Capteurs de type Ultrasons, Magnetiques et Electrostatiques," *Instrum., Mesure, Metrol.-Evalu. Nondestr.* (in French), **1**, 101–125.
- Schmerr, L. W. (1998). *Fundamentals of Ultrasonic Nondestructive Evaluation-A Modeling Approach* (Plenum, New York).
- Wooh, S. C., and Shi, Y. (1999). "A design strategy for phased arrays," in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York), Vol. 18A, pp. 1061–1068.

Lamb mode conversion at edges. A hybrid boundary element–finite-element solution

José M. Galán^{a)} and Ramón Abascal^{b)}

Escuela Superior de Ingenieros, Camino de los Descubrimientos s/n, E-41092, Sevilla, Spain

(Received 1 October 2000; revised 1 July 2002; accepted 3 December 2004)

Two general and flexible numerical techniques based on the finite-element and boundary element methods developed by the authors in a previous paper are applied to study Lamb wave propagation in multilayered plates and Lamb mode conversion at free edges for frequencies beyond the first cutoff frequency. Both techniques are supported by a meshing criterion which guarantees the accuracy of the results when a condition is fulfilled. A finite-element formulation is directly applicable to study Lamb wave propagation and reflection by simple obstacles such as a flat edge. In order to tackle Lamb wave diffraction problems by defects with more complex geometries, a hybrid boundary element–finite-element formulation is used. This technique provides a major improvement with respect to the only previous boundary element application on Lamb waves: the connecting boundary might be placed as close to the reflector as desired, reducing greatly the requirement on mesh size. Two main application problems on practical metallic plates are studied and compared with reported numerical, theoretical, and experimental results: (1) Lamb wave propagation in degraded titanium diffusion bonds, and (2) Lamb mode conversion at inclined or perpendicular free edges of steel plates for frequencies beyond the first cutoff frequency. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1857525]

PACS numbers: 43.20.Gp, 43.20.Bi, 43.35.Cg [VWS]

Pages: 1777–1784

I. INTRODUCTION

The use of elastic guided waves for nondestructive testing (NDT) purposes is developing rapidly.¹ Their long propagation distance and their ability to travel along the structures are advantages with respect to bulk wave ultrasonic techniques. Some successful applications have already been reported.^{2,3} This paper focuses on Lamb waves, which are guided waves in plates. A thorough theoretical analysis of these waves can be found in the literature.^{4–6}

A further development of Lamb wave NDT techniques requires understanding the interaction with defects and geometrical features. Lamb wave scattering is more complex than bulk wave scattering because the whole set of propagating modes is generated in the scattered field. This behavior restrains an analytical solution to very simple cases, and a numerical approach is therefore required.

The reflection at the edge of a semi-infinite plate is the simplest case of Lamb wave scattering, but it provides insight into its essential features. The first studies were aimed to explain the edge resonance phenomena.^{7–12} Later, the mode conversion for frequencies above the first cutoff frequency was studied theoretically,^{11,8} numerically,¹³ and experimentally.¹⁴ Lamb wave scattering by more general and realistic obstacles and defects (such as cracks, delaminations, or corrosion) has also been studied by other authors. Some theoretical approaches^{15–18} are available which express the field in the plate as an analytical normal mode expansion, but they are valid for very particular geometries. In the numerical studies, the finite-element (FE) method is the most

widely used technique. Two different implementations are considered: (1) Frequency-domain analysis, where a normal mode expansion of the far field is connected to a FE discretization of the defect neighborhood,^{19–23} (2) Time-domain analysis, where the scattering of ultrasonic pulses of finite duration is considered.^{24–26} Alleyne and Cawley²⁴ perform a postprocessing of the results to extract the modal content of the signals through a 2D-FFT. Verdict *et al.*²⁵ and Zgonc and Achenbach²⁶ use the time records directly to train neural networks for defect detection and characterization.

The boundary element (BE) method has not been used in guided wave scattering problems until very recently.^{13,27} The pioneering works by Cho and Rose combine the 2D frequency-domain BE method with a truncated analytical normal mode expansion including propagating modes only. This hybrid method was applied to study the reflection at free edges of semi-infinite steel plates,¹³ and the scattering by elliptical and rectangular surface defects on steel and aluminum plates.²⁷

The authors presented in a previous paper²⁸ two numerical techniques for the study of Lamb waves in plates which simplify the calculation of dispersion curves and diffraction by defects. The first technique is an FE formulation^{29–31} with two important features: (a) dispersion curves can be approximated easily and accurately up to the maximum frequency of interest by solving a quadratic eigensystem; (b) the FE dynamic stiffness matrix of the semi-infinite plate is an accurate approximation of an absorbing boundary condition for Lamb waves, which automatically takes into account the radiation of energy towards infinity in the form of propagating modes and includes the near-field effect of the evanescent modes. The second technique is a hybrid BE–FE formulation which uses the FE dynamic stiffness matrix of the semi-

^{a)}Electronic mail: mferman@us.es

^{b)}Author to whom correspondence should be addressed. Electronic mail: abascal@us.es

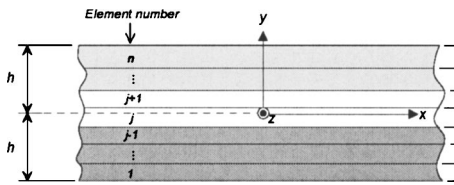


FIG. 1. FE discretization of a multilayered plate.

infinite plate as an absorbing boundary condition on a BE region modeling the vicinity of a defect. The coupling between the FE and BE regions might be placed as close to the defect as desired. Both formulations are complemented by a meshing criterion which assures the accuracy of the results.

Additional results on Lamb wave propagation and scattering in metallic materials are presented in this paper. A multilayered plate which models a degraded titanium diffusion bond is studied to check the performance and applicability of the FE technique to nonhomogeneous plates. The reflection of high Lamb modes at the edge of a semi-infinite steel plate is analyzed. Reflection of Lamb waves at inclined edges is solved by the hybrid BE–FE technique.

II. AN FE FORMULATION FOR LAMB WAVES

A. Dynamic stiffness matrix of an infinite plate

Let us consider an infinite multilayered plate of thickness $2h$ under plane strain as shown in Fig. 1. Displacements corresponding to plane harmonic waves propagating in the x direction are considered. The plate cross section is discretized with n elements and N nodes, and shape functions are used to interpolate the variables. One-dimensional isoparametric (linear or quadratic) finite elements are used. Each element must be homogeneous, although the mesh may have elements of different materials.

The global stiffness and mass matrices for the infinite plate can be obtained and assembled into the global system of equations, rendering²⁸

$$(k^2\mathbf{A} + ik\mathbf{B} + \mathbf{G} - \omega^2\mathbf{M})\mathbf{d} = \mathbf{f}, \quad (1)$$

where the matrices \mathbf{A} , \mathbf{G} , and \mathbf{M} are symmetric and \mathbf{B} is antisymmetric, and \mathbf{d} and \mathbf{f} are vectors containing the nodal values of the displacements and consistent forces, respectively. The homogeneous problem with $\mathbf{f} = \mathbf{0}$ is used to obtain the Lamb wave modes and wave numbers.

B. Dynamic stiffness matrix of a semi-infinite plate

A semi-infinite plate of thickness $2h$ extending over the range $x \in (-\infty, x_0]$ is considered (see Fig. 2). The elastodynamic field can be expressed as a superposition of Lamb modes (normal mode expansion^{7,11}) where only modes satis-

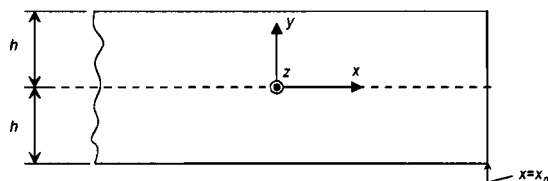


FIG. 2. Geometry of a semi-infinite plate.

fying the radiation condition are included. These are nonevanescent modes propagating towards $x = -\infty$ or evanescent modes with decaying amplitude for $x \rightarrow -\infty$. Taking into account that the dependence with x and t is $\exp[i(kx + \omega t)]$, the former modes have a real, positive wave number (except for backward waves) and the latter have a complex wave number with negative imaginary part.

For a certain frequency ω , the FE formulation approximates the Lamb modes with the solutions of the quadratic eigensystem in Eq. (1). A mesh of N nodes produces $4N$ eigenvalues for 2D problems, from which only $2N$ satisfy the radiation condition. These $2N$ approximated modes, which include propagating and evanescent modes, are used in a normal mode expansion of the reflected field at the edge, providing a relationship between displacements and consistent nodal forces at section $x = x_0$:²⁸ $\mathbf{F}^{\text{refl}} = \mathbf{S}\mathbf{u}^{\text{refl}}$, where \mathbf{S} is the FE dynamic stiffness matrix of the semi-infinite plate which is independent of x_0 .

1. Incident wave

Some excitation at $x < x_0$ not included in the previous model can generate a Lamb wave which will propagate through the plate towards the edge, $x = x_0$. Superposition can be applied by expressing the total field as the sum of an incident and a reflected wave: $\mathbf{u} = \mathbf{u}^{\text{I}} + \mathbf{u}^{\text{refl}}$ and $\mathbf{F} = \mathbf{F}^{\text{I}} + \mathbf{F}^{\text{refl}}$, where the incident field is known. A relationship between displacements and consistent forces in the total field can be written: $\mathbf{F} - \mathbf{F}^{\text{I}} = \mathbf{S}(\mathbf{u} - \mathbf{u}^{\text{I}})$.

The excitation source can generate all the possible modes (propagating and evanescent) at that frequency. If the source is sufficiently far from the edge, the evanescent modes have zero amplitude at section $x = x_0$ because their amplitude decays exponentially with distance. Only propagating modes will reach the edge. In this paper, incident plane harmonic Lamb waves containing a single propagating mode are considered. By taking into account the boundary condition at $x = x_0$, participation factors of modes in the reflected field are calculated.

2. Partition of energy flux. Conservation of energy

The total reflected energy flux is computed from the FE results as follows:

$$E_{\text{ref}} = \sum_{\text{propagating } m} |\alpha^{(m)}|^2 E^{(m)}, \quad (2)$$

where $\alpha^{(m)}$ is the participation coefficient of mode m and $E^{(m)}$ is the energy flux averaged over a period carried by the propagating mode m . The summation is extended over the propagating modes only, since the evanescent modes do not carry any net energy flux when averaged over a period.

The principle of conservation of energy applied to a semi-infinite plate states that the energy flux transported by the incident wave E_{inc} must be equal to the energy flux transported by the reflected wave E_{ref} through any cross section of the plate. In this work two main sources of approximation are considered: an FE discretization is used and only a finite number of modes is retained in the normal mode expansion. They both introduce an error in the solution which causes the

violation of the energy-flux conservation principle. A first estimation of the accuracy of the results is that the energy-flux conservation is satisfied within a small margin, which is a necessary but not sufficient condition to validate the numerical results. All the numerical results for Lamb wave scattering in semi-infinite plates presented subsequently in this paper satisfy energy-flux conservation within a 1% margin, which is a very tight limit in comparison to previous published work^{20,13} where even a 10% error is allowed.

C. Meshing criterion

An analysis of the error when the FE stiffness matrix is used instead of the exact stiffness matrix for a thin, homogeneous plate was carried out by the authors in a previous paper,²⁸ from which a meshing criterion was proposed that can be applied to calculate dispersion curves and to solve Lamb wave scattering problems. In order to obtain accurate results, the meshing criterion requires the fulfillment of the following condition:²⁸

$$\frac{\lambda_T}{L} > \beta, \quad (3)$$

where L is the element size, $\lambda_T = 2\pi c_T/\omega$ is the wavelength of transverse waves at frequency ω , $\beta=20$ for linear elements, and $\beta=4$ for quadratic elements. This criterion has a clear physical interpretation: in order to get accurate results with an FE mesh, the number of elements per transverse wavelength must be greater than β . It agrees with the usual rule-of-thumb of FE in dynamics. From the viewpoint of a practical application, the following two solutions are provided by the meshing criterion:

- (i) Determination of element size L to get accurate results up to a given maximum frequency ω_{\max} , which yields $L < 2\pi c_T/\beta\omega_{\max}$;
- (ii) Determination of the maximum frequency that a given mesh of element size L can solve accurately: $\omega < 2\pi c_T/\beta L$.

III. SOME FE RESULTS

A. Numerical calculation of dispersion curves

The FE formulation is able to produce dispersion curves for Lamb waves in a simple manner by solving the eigenproblem in Eq. (1). In this paper, dispersion curves are calculated by means of a wave number sweep using IMSL eigenroutines.³²

1. Homogeneous isotropic plate

The dispersion curves for Lamb waves in homogeneous plates were calculated by Mindlin and other authors in the 50's,⁴⁻⁶ by numerically tracking the roots of the Rayleigh-Lamb frequency equations. Although the dimensionless dispersion curves only depend on Poisson's ratio,⁵ the numerical calculation requires complete definition of the material and dimensions of the plate. A plate of thickness $2h = 2$ mm with Poisson's ratio $\nu=0.25$ and bulk wave velocities $c_L = 1$ km/s and $c_T = 1/\sqrt{3}$ km/s is considered.

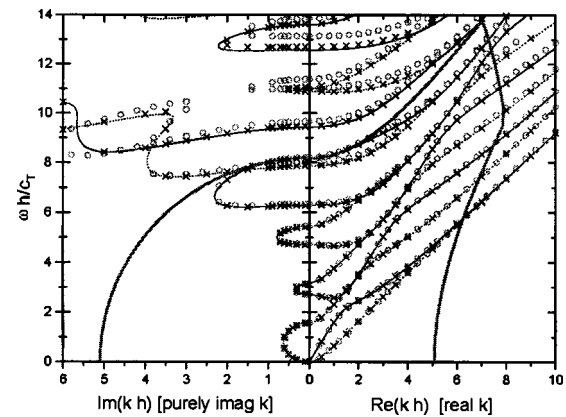


FIG. 3. Dispersion curves with 24 linear elements and 12 quadratic elements.

The dimensionless FE results obtained are plotted and compared with exact dispersion curves in Fig. 3. Exact branches corresponding to symmetric modes are plotted with a thin continuous line, exact branches corresponding to anti-symmetric modes with a thin dotted line, FE results with linear elements with \circ dots and numerical results with quadratic elements with \times crosses. The meshes of linear and quadratic elements used have the same number of d.o.f. The agreement is excellent up to the limit defined by the meshing criterion. The thick black line encloses an area where the error in the linear element stiffness matrices is below 5%.

2. Titanium diffusion bond

Diffusion-bonded members are used in aerospace structures. The quality of the bond is an aspect which must be assessed by some NDT technique, among which Lamb waves have been considered.¹ The thin diffusion bond interface can be modeled as a thin layer having distinct properties from the titanium (Ti). Bond quality can be represented in this manner by assigning degraded properties to the bond layer. In this paper the FE formulation is applied to calculate dispersion curves for a titanium diffusion bond under two different bonding conditions: (1) perfect bonding, where the bond layer has the same properties as Ti and therefore a homogeneous Ti plate is considered; (2) poor-quality bonding, where the bond layer has degraded properties compared to Ti. A nonuniform mesh of 27 quadratic elements is used, which is guaranteed by the meshing criterion to provide accurate results for frequencies up to $\omega h/(c_T)_{Ti} < 15.8$. The geometry is perfectly symmetrical with respect to the plate midplane, and the dimensions are total thickness $2h = 2.01$ mm, bond layer thickness 0.01 mm. The material properties used in the calculations are for titanium $\rho = 4.46$ kg/dm³, $c_L = 6.06$ km/s, $c_T = 3.23$ km/s and for the poor-quality bond layer $\rho = 2.2$ kg/dm³, $c_L = 1.35$ km/s, $c_T = 0.55$ km/s. Numerical results are plotted in Fig. 4, where they are compared with the exact results for the homogeneous Ti plate. One observes a shift to the right in the $k-\omega$ dispersion curves of the degraded bond with respect to the perfect bond case, which agrees with experimental observations and previous numerical calculations.¹

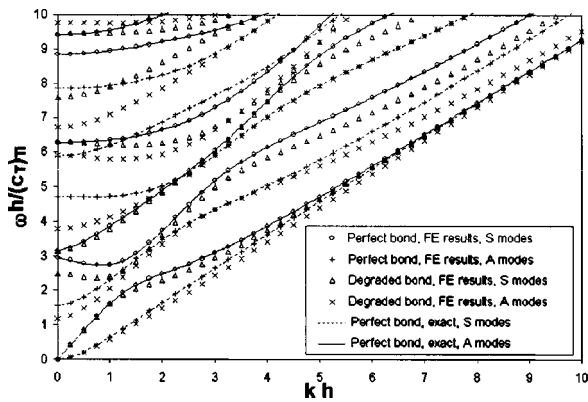


FIG. 4. Dispersion curves for titanium diffusion bond.

Dispersion curves for the imperfect bond are compared in Fig. 5 with those obtained with DISPERSE,³³ a commercial software to generate dispersion curves which makes use of the global matrix method. A complete agreement between both results is observed.

B. Reflection of an S0 mode by a free end

The reflection of an S0 mode by a free edge of a semi-infinite homogeneous plate was solved by Gazis and Mindlin⁷ for frequencies below the first cutoff frequency, by Torvik¹¹ for frequencies below the third cutoff frequency, and by Gregory and Gladwell⁸ for even higher frequencies.

A semi-infinite homogeneous plate of thickness $2h$ with a free end at $x=0$ is considered. The material is linear, elastic, and isotropic, with a Poisson's ratio $\nu=0.25$. For numerical calculations, a plate thickness $2h=2$ mm, and bulk wave velocities $c_L=1$ km/s and $c_T=1/\sqrt{3}$ km/s are considered.

Since the reflector (free end) is symmetric, there cannot be any mode conversion between symmetric and antisymmetric modes. The numerical results reproduce this effect exactly. The minimum number of elements required by the meshing criterion in order to accurately reproduce Fig. 4 from Ref. 8 is $n_{\min}=11$. Numerical results with 11, 22, and 44 quadratic elements are represented and compared with those of Gregory and Gladwell in Fig. 6, showing excellent agreement.

Shaw,¹⁰ while studying experimentally the axisymmetric resonant vibrations of thick barium titanate disks ($\nu=0.31$),

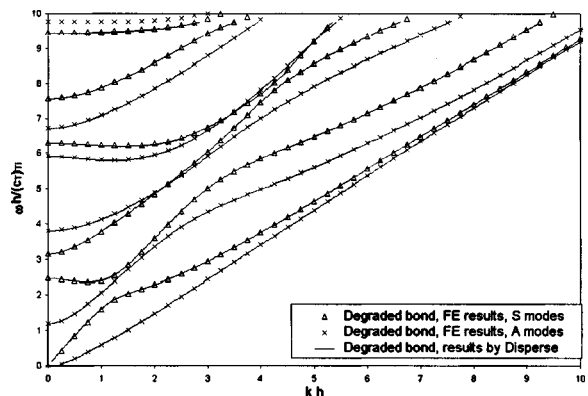


FIG. 5. Dispersion curves for imperfect diffusion bond. Comparison with DISPERSE.

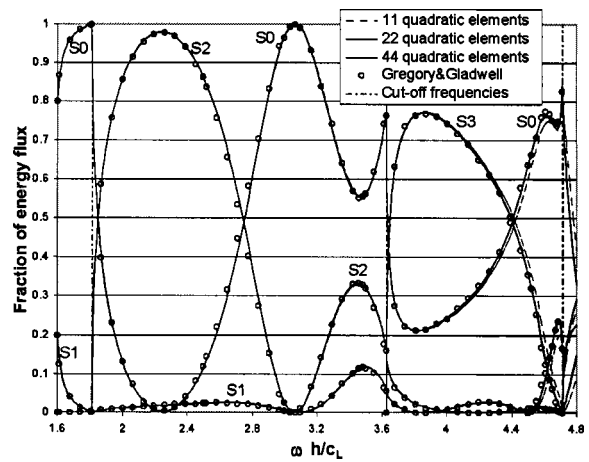


FIG. 6. Reflection of mode S0 by a free end. Results with quadratic elements.

observed the existence of an edge resonance below the first cutoff frequency which he could not explain theoretically. A theoretical explanation was first given by Gazis and Mindlin,⁷ who showed that it is produced by a pair of evanescent modes corresponding to complex conjugate roots of the frequency equation when the S0 mode impinges on the edge. A similar edge resonance was found at the free edges of semi-infinite plates. Later, Torvik,¹¹ Tsao and Auld,¹² and Gregory and Gladwell⁸ confirmed those conclusions by means of more refined calculations. Using the FE formulation, a resonance in the second and third symmetric modes was found at a frequency $(\omega h/c_L)^2=1.493$, where they are evanescent, which agrees with the results reported by those authors.

C. Reflection of higher Lamb modes by a free end

The reflection of higher Lamb modes by a free edge of a semi-infinite homogeneous plate was studied experimentally using photoelasticity by Zhang *et al.*,¹⁴ and numerically by Cho and Rose,¹³ using a combination of a normal mode expansion technique and BEM with piecewise constant elements. The approach of the latter paper is based on expanding the scattered field as a superposition of Lamb modes (normal mode expansion) and keeping only propagating modes. The dispersion curves are obtained by numerically finding the roots of the exact Rayleigh–Lamb frequency

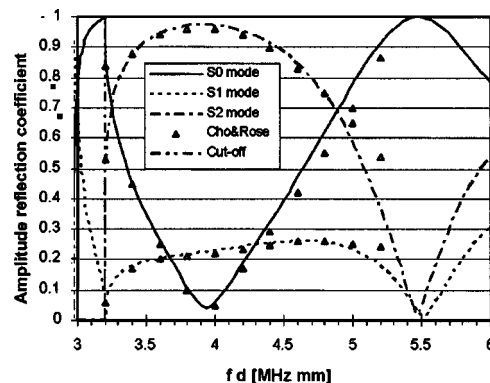


FIG. 7. Reflection of S0 mode at a flat free edge.

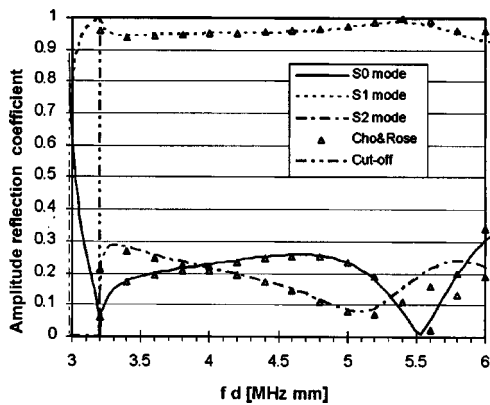


FIG. 8. Reflection of $S1$ mode at a flat free edge.

equation for the plate. The analytical expressions for displacements and stresses are used in order to obtain a non-reflecting boundary condition. Since only propagating modes are included in the modal superposition, the absorbing boundary must be placed far enough from the edge (at a distance longer than one wavelength of the incident mode) in order to avoid the near-field effect of evanescent modes. This approximation is satisfactory except in the neighborhood of the cutoff frequencies, where evanescent modes with wave numbers having very small imaginary part and therefore with a very weak exponential decay are found.

The FE formulation is used to study the reflection of higher modes at free flat edges. In order to compare with Cho and Rose's results, a semi-infinite homogeneous plate of thickness $d=2h=1$ mm with a free end at $x=0$ is considered. The material is common steel, with density $\rho=7.8$ kg/dm³ and bulk wave velocities $c_L=5.94$ km/s and $c_T=3.2$ km/s. Since the reflector (free edge) is symmetric, no mode conversion between symmetric and antisymmetric modes can occur. The numerical results reproduce this effect exactly.

The maximum frequency considered in Cho and Rose's paper¹³ is $f=6$ MHz. The minimum number of quadratic elements required by the meshing criterion for frequencies up to that value is $n_{\min}=8$. A uniform FE mesh of 8 quadratic elements was used to get the numerical results which are plotted in Figs. 7-11. Using a uniform mesh of 16 elements does not change the results practically. The agreement with Cho and Rose, who used more than 200 constant elements, is

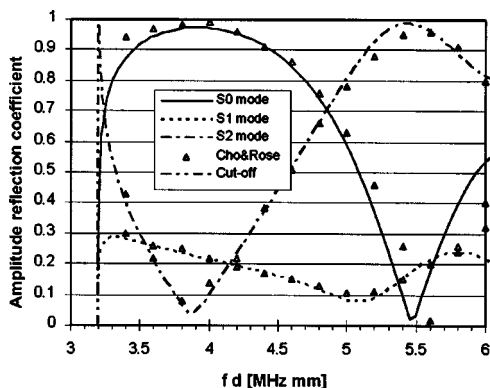


FIG. 9. Reflection of $S2$ mode at a flat free edge.

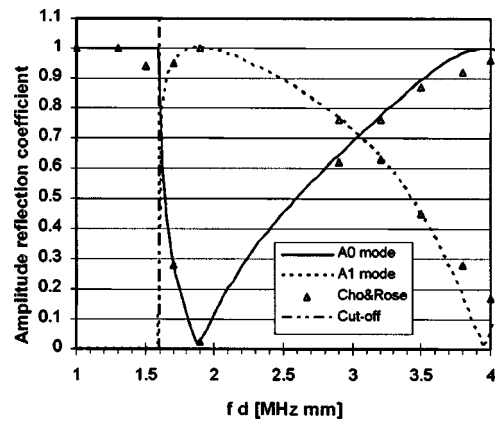


FIG. 10. Reflection of $A0$ mode at a flat free edge.

good. Moreover, since the formulation used in this work retains the evanescent modes in the far-field expansion, the results are also accurate in the neighborhood of the cutoff frequencies. Cho and Rose reported some values of the reflected amplitudes above 1 (see Fig. 11), which clearly violates the energy-conservation principle, but never beyond the $\pm 10\%$ bound allowed in their calculations.

The FE results presented satisfy the energy-flux reciprocity principle, i.e., when a p mode is incident on to the edge, the fraction of energy which is transformed into an n mode is precisely the same as that transformed into a p mode upon incidence of an n mode onto the edge.

IV. THE BE METHOD

The BE method (BEM) is a numerical method that has been widely and successfully used to solve elastodynamic problems.^{34,35} Some of its advantages with respect to other domain-type methods, such as FE and finite difference methods, are (1) reduction of the dimension of the problem, since only the surface of the domain has to be discretized; (2) better performance in the analysis of high-stress gradient problems, such as cracks. In this section the formulation of BEM is briefly described.

The reciprocal theorem for two time-harmonic elastodynamic states of the same frequency ω on a body Ω with surface Γ can be written as³⁴

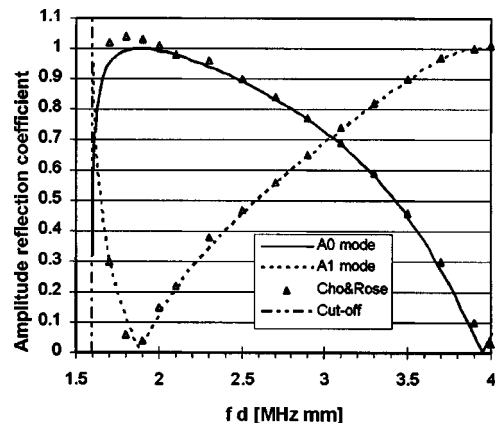


FIG. 11. Reflection of $A1$ mode at a flat free edge.

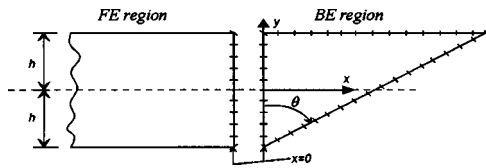


FIG. 12. Hybrid BE–FE model for the inclined edge.

$$\int_{\Gamma} t_i u_i^* d\Gamma + \int_{\Omega} \rho b_i u_i^* d\Omega = \int_{\Gamma} t_i^* u_i d\Gamma + \int_{\Omega} \rho b_i^* u_i d\Omega, \quad (4)$$

where the field variables u_i , t_i , b_i , u_i^* , t_i^* , and b_i^* are the components of the displacements, tractions, and body forces, respectively, that depend on position \mathbf{x} and frequency ω and might be real or complex-valued.

The elastodynamic fundamental solution corresponds to the displacements and stresses produced by a time-dependent point load in an unbounded medium. It was first presented by Stokes.⁵ Assuming that the second state in Eq. (4) (the one denoted by*) is the fundamental solution for a unit time-harmonic load of frequency ω applied along the l direction at a point $\mathbf{x}=\xi \in \Omega$, taking the collocation point ξ to the boundary by a limiting process, and extracting the singularity at $r=0$ due to the fundamental solution, renders

$$c_{lk}(\xi) u_k(\xi, \omega) + \int_{\Gamma} t_{lk}^*(\mathbf{x}, \xi, \omega) u_k(\mathbf{x}, \omega) d\Gamma(\mathbf{x}) = \int_{\Gamma} u_{lk}^*(\mathbf{x}, \xi, \omega) t_k(\mathbf{x}, \omega) d\Gamma(\mathbf{x}), \quad (5)$$

where the integrals are Cauchy principal value integrals, u_k and p_k are the boundary displacements and tractions, and the matrix $c_{lk}(\xi)$ only depends on the smoothness and the local geometry of the boundary at the collocation point ξ .

The numerical solution of Eq. (5) requires the discretization of the boundary Γ into elements $\Gamma = \sum_{j=1}^n \Gamma_j$, where n is the number of elements and Γ_j the boundary of element j . Over each element, geometry, displacements, and tractions are interpolated in terms of their nodal values by means of shape functions. In BEM, displacements and tractions are interpolated independently, in contrast with FEM where stresses (and tractions) are obtained as derivatives of the interpolated displacements. Substituting the element discretization into Eq. (5) yields two equations relating the displacement of the collocation point with the nodal displacements and tractions of all boundary nodes. A system of equations in terms of the nodal values as the only unknowns can be obtained by writing that equation at all boundary nodes, which after taking the boundary conditions into account provides the solution on the boundary.

V. HYBRID BE–FE FORMULATION

The problem of Lamb wave interaction with complex obstacles and defects is tackled by dividing the geometry in two regions, as shown in Fig. 12: (1) neighborhood of the defect and (2) far field. The former region is modeled with BE, which is proven very accurate and successful in the study of harmonic problems and can reproduce accurately

many kinds of obstacles, such as inclusions, voids, delaminations, and cracks. The latter region is modeled with FE by means of a truncated normal modal expansion which gives rise to the dynamic stiffness matrix of the semi-infinite plate \mathbf{S} . This matrix represents a nonreflecting boundary condition for Lamb waves with a very high accuracy up to a frequency limit defined by the meshing criterion [Eq. (3)]. It reproduces in a precise manner the energy flux towards infinity through a semi-infinite plate in the form of propagating Lamb modes. The near-field effect of evanescent modes is also taken into account.

This modeling strategy is similar to that followed by Cho and Rose,^{13,27} but our approach introduces three main improvements.

- (1) Not only propagating but also evanescent modes are used in the far-field expansion. Therefore, the response in the neighborhood of the cutoff frequencies is reproduced better;
- (2) The dispersion curves and modes are obtained numerically by the FE method, which greatly simplifies the calculation of the wave numbers and permits the extension to sandwich plates without any additional effort; and
- (3) Quadratic boundary elements are used instead of constant elements.

This hybrid BE–FE formulation requires a coupling between the BE and FE regions. Several techniques can be found in the literature.^{36,37} In our approach, compatibility and equilibrium are enforced variationally and at an element level.²⁸

It was shown²⁸ that the fictitious absorbing boundary can be placed as close to the reflector as desired without any decrease in accuracy. This is in clear contrast with Ref. 13, where the coupling between normal mode expansion and BE region had to be done at a distance d_f longer than an incident wavelength. This distance becomes large for high frequencies and consequently demands a large number of elements. The explanation for this behavior is that Ref. 13 does not include evanescent modes in the normal mode expansion, which are essential in the vicinity of reflectors. The number of elements required to solve the problem with the proposed hybrid formulation is therefore greatly reduced.

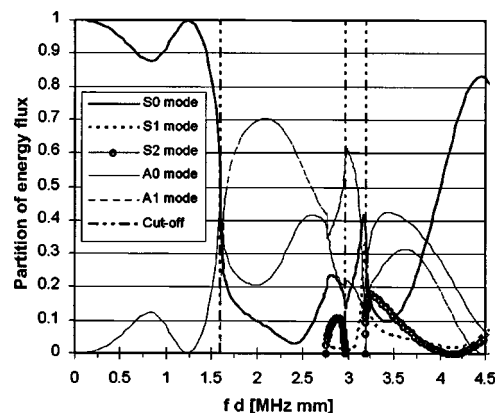


FIG. 13. Reflection of S0 mode at a 45 deg inclined free edge.

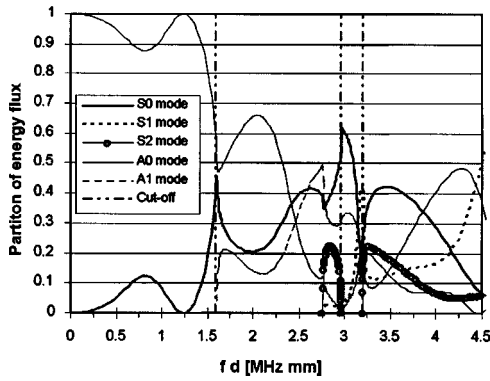


FIG. 14. Reflection of A0 mode at a 45 deg inclined free edge.

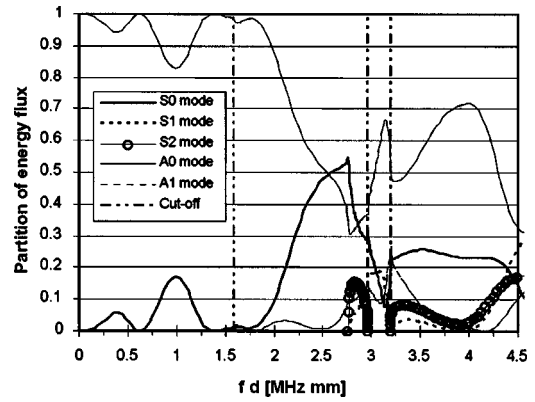


FIG. 16. Reflection of A0 mode at a 60 deg inclined free edge.

VI. SOME HYBRID BE-FE RESULTS

A. Reflection of S0 and A0 modes by inclined free edges

Let us consider a semi-infinite homogeneous plate of thickness $2h = 1$ mm with an inclined free edge with an inclination angle θ with respect to the vertical axis. A Lamb mode (S0 or A0) is incident from $x = -\infty$. The material is common steel, with density $\rho = 7.8$ kg/dm³ and bulk wave velocities $c_L = 5.94$ km/s and $c_T = 3.2$ km/s. Because of the geometry this problem cannot be solved using only the FE formulation of Sec. II B. The hybrid BE-FE model for this problem is shown in Fig. 12. The fictitious boundary is placed as close to the edge as possible in order to reduce the number of elements in the BE region. Quadratic elements of size $L = 0.125$ mm are used to mesh both regions. According to the meshing criterion, this element size gives accurate results up to 6 MHz.

This problem is solved for two different angles $\theta = 60$ deg and 45 deg using meshes of 38 and 28 quadratic elements, respectively. The results of the partition of energy flux are shown in Figs. 13-16. Since the reflector is not symmetric, a mode conversion between symmetric and antisymmetric modes occurs and can be observed in the figures. The numerical results satisfy the energy-flux reciprocity principle, as shown in Fig. 17.

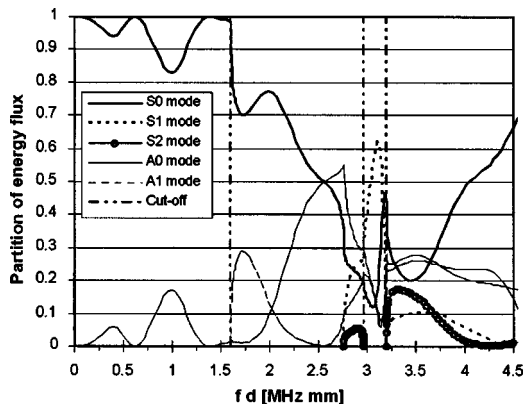


FIG. 15. Reflection of S0 mode at a 60 deg inclined free edge.

VII. CONCLUSIONS

Two numerical techniques based on the finite-element- and boundary element methods developed by the authors in a previous paper are applied to practical cases of Lamb waves in metallic plates with excellent results.

Dispersion curves for degraded titanium diffusion bonds, modeled as nonhomogeneous three-layer plates, have been calculated and compared with previously reported experimental and numerical results. A shift of the degraded curves with respect to the perfect bond curves is observed, in complete agreement with previous work.

Reflection of high Lamb modes at the free edge of a semi-infinite steel plate are studied and compared with previous work by Cho and Rose. The agreement is good, with a much smaller mesh and computational effort. Reflection at an inclined edge of a semi-infinite steel plate is studied using the hybrid BE-FE technique for two different angles, $\theta = 60$ deg and 45 deg.

The meshing criterion proposed by the authors has been used in all the examples and is shown to provide meshes whose results are accurate up to the desired frequency. The accuracy of the scattering results is such that energy-flux conservation and reciprocity are satisfied within a 1% margin.

The hybrid technique developed by the authors contains

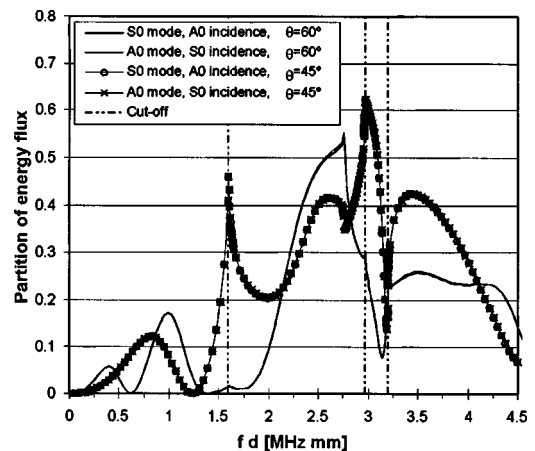


FIG. 17. Check of the reciprocity principle for Lamb mode reflection at inclined free edges.

two main improvements as compared to previous BE techniques for Lamb wave problems.

- (1) The fictitious boundary might be placed as close to the reflector as desired, since evanescent modes are included in the normal mode expansion. Hence, the mesh size requirement is greatly reduced.
- (2) The accuracy of the numerical results is under control by the proposed meshing criterion.

In conclusion, the two numerical techniques presented by the authors for the study of Lamb waves in homogeneous or sandwich plates greatly simplify the calculation of dispersion curves and diffraction by defects. They provide a flexible and general tool to solve Lamb wave scattering problems that can be readily extended to study defects of more complex geometry in semi-infinite and infinite plates.

ACKNOWLEDGMENTS

This work was funded by Spain's DGICYT in the framework of Project DPI 2000-1642.

¹J. L. Rose, *Ultrasonic Waves in Solid Media* (Cambridge University Press, Cambridge, England, 1999).

²J. L. Rose, K. M. Rajana, and F. T. Carr, "Ultrasonic guided wave inspection concepts for steam generator tubing," *Mater. Eval.* **52**(2), 307–316 (1994).

³J. L. Rose, D. Jiao, and J. Spanner, Jr., "Ultrasonic guided wave NDE for piping," *Mater. Eval.* **54**(11), 1310–1313 (1996).

⁴I. A. Viktorov, *Rayleigh and Lamb Waves* (Plenum, New York, 1967).

⁵J. D. Achenbach, *Wave Propagation in Elastic Solids* (North-Holland, New York, 1973).

⁶K. F. Graff, *Wave Motion in Elastic Solids* (Clarendon, Oxford, 1975).

⁷D. C. Gazis and R. D. Mindlin, "Extensional vibrations and waves in a circular disk and a semi-infinite plate," *J. Appl. Mech.* **27**, 541–547 (1960).

⁸R. D. Gregory and I. Gladwell, "The reflection of a symmetric Rayleigh–Lamb wave at the fixed or free edge of a plate," *J. Elast.* **13**, 185–206 (1983).

⁹M. Koshihba, S. Karakida, and M. Suzuki, "Finite-element analysis of edge resonance in a semi-infinite elastic plate," *Electron. Lett.* **19**, 256–257 (1983).

¹⁰E. A. G. Shaw, "On the resonant vibrations of thick barium titanate disks," *J. Acoust. Soc. Am.* **28**, 38–50 (1956).

¹¹P. J. Torvik, "Reflection of wave trains in semi-infinite plates," *J. Acoust. Soc. Am.* **41**(2), 346–353 (1967).

¹²B. A. Auld and M. T. Tsao, "A variational analysis of edge resonance in a semi-infinite plate," *IEEE Trans. Sonics Ultrason.* **SU-24**(5), 317–326 (1977).

¹³Y. Cho and J. L. Rose, "A boundary element solution for a mode conversion study on the edge reflection of Lamb waves," *J. Acoust. Soc. Am.* **99**(4), 2097–2109 (1996).

¹⁴S. Y. Zhang, J. Z. Shen, and C. F. Ying, "The reflection of the Lamb wave by a free plate edge: Visualization and theory," *Mater. Eval.* **46**, 638–641 (1988).

¹⁵S. I. Rokhlin, "Interaction of Lamb waves with elongated delaminations in thin sheets," *Int. Adv. Nondestr. Test.* **6**, 263–285 (1979).

¹⁶S. I. Rokhlin, "Diffraction of Lamb waves by a finite crack in an elastic layer," *J. Acoust. Soc. Am.* **67**(4), 1157–1165 (1980).

¹⁷M. Tan and B. A. Auld, "Normal mode variational method for two and three dimensional acoustic scattering in an isotropic plate," *IEEE Ultrasonics Symposium*, 1980, pp. 857–861.

¹⁸S. I. Rokhlin, "Lamb wave interaction with lap-shear adhesive joints: Theory and experiment," *J. Acoust. Soc. Am.* **89**(6), 2758–2765 (1991).

¹⁹M. Koshihba, S. Karakida, and M. Suzuki, "Finite-element analysis of Lamb wave scattering in an elastic plate waveguide," *IEEE Trans. Sonics Ultrason.* **SU-31**(1), 18–25 (1984).

²⁰Y. N. Al-Nassar, S. K. Datta, and A. H. Shah, "Scattering of Lamb waves by a normal rectangular strip weldment," *Ultrasonics* **29**, 125–132 (1991).

²¹S. K. Datta, Y. Al-Nassar, and A. H. Shah, "Lamb wave scattering by a surface-breaking crack in a plate," in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by O. D. Thompson and D. E. Chimenti (Plenum, New York, 1991), Vol. 10, pp. 97–104.

²²M. R. Karim, M. A. Awal, and T. Kundu, "Elastic wave scattering by cracks and inclusions in plates: In-plane case," *Int. J. Solids Struct.* **29**(19), 2355–2367 (1992).

²³A. Mal, Z. Chang, and M. Gorman, "Interaction of Lamb waves with defects in a semi-infinite plate," in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by O. D. Thompson and D. E. Chimenti (Plenum, New York, 1997), Vol. 16, pp. 153–160.

²⁴D. N. Alleyne and P. Cawley, "The interaction of Lamb waves with defects," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **39**(3), 381–397 (1992).

²⁵G. S. Verdict, P. H. Gien, and C. P. Burger, "Finite element study of Lamb wave interactions with holes and through thickness defects in thin metal plates," in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by O. D. Thompson and D. E. Chimenti (Plenum, New York, 1992), Vol. 11, pp. 97–104.

²⁶K. Zgonc and J. D. Achenbach, "A neural network for crack sizing trained by finite element calculations," *NDT & E Int.* **29**, 2097–2109 (1996).

²⁷Y. Cho, D. D. Hongerholt, and J. L. Rose, "Lamb wave scattering analysis for reflector characterization," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **44**(1), 44–52 (1997).

²⁸J. M. Galan and R. Abascal, "Numerical simulation of Lamb wave scattering in semi-infinite plates," *Int. J. Numer. Methods Eng.* **53**, 1145–1173 (2002).

²⁹V. Lotfi, "Analysis of the response of dams to earthquakes," Technical Report GR86-2, CE Dept., University of Texas at Austin, 1986.

³⁰E. Kausel, "Forced vibrations of circular foundations on layered media," Technical Report R74-11, Department of Civil Engineering, MIT, 1974.

³¹G. R. Liu and J. D. Achenbach, "A strip element method to analyze wave scattering by cracks in anisotropic laminated plates," *J. Appl. Mech.* **62**, 607–613 (1995).

³²IMSL. *IMSL MATH/LIBRARY User's Manual, Version 3.0* (IMSL, Houston, 1994).

³³B., Pavlakovic, M. J. S., Lowe, D. N. Alleyne, and P. Cawley, "DISPERSE: A general purpose program for creating dispersion curves," in *Review of Progress in Quantitative NDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1997), Vol. 16, pp. 185–192.

³⁴J. Dominguez, *Boundary Elements in Dynamics* (CMP, Southampton, 1993).

³⁵J. Dominguez and E. Alarcon, *Progress in Boundary Element Method* (Pentech, London, 1981), Chap. 7.

³⁶T. Belytschko and Y. Y. Lu, "A variationally coupled FE–BE method for transient problems," *Int. J. Numer. Methods Eng.* **37**, 91–105 (1994).

³⁷F. Guarracino, V. Minutolo, and L. Nunziante, "A simple analysis of soil–structure interaction by BEM–FEM coupling," *Eng. Anal. Boundary Elem.* **10**, 280–290 (1992).

Dispersion of waves in porous cylinders with patchy saturation: Formulation and torsional waves

James G. Berryman^{a)}

University of California, Lawrence Livermore National Laboratory, P.O. Box 808 L-200,
Livermore, California 94551-9900

Steven R. Pride^{b)}

Lawrence Berkeley National Laboratory, Earth Sciences Division, 1 Cyclotron Road MS 90-1116,
Berkeley, California 94720

(Received 9 September 2004; revised 23 December 2004; accepted 31 December 2004)

Laboratory experiments on wave propagation through saturated and partially saturated porous media have often been conducted on porous cylinders that were initially fully saturated and then allowed to dry while continuing to acquire data on the wave behavior. Since it is known that drying typically progresses from outside to inside, a sensible physical model of this process is concentric cylinders having different saturation levels—the simplest example being a fully dry outer cylindrical shell together with a fully wet inner cylinder. We use this model to formulate the equations for wave dispersion in porous cylinders for patchy saturation (i.e., drainage) conditions. In addition to multiple modes of propagation obtained numerically from these dispersion relations, we find two distinct analytical expressions for torsional wave modes. We solve the resulting torsional wave dispersion relation for two examples: Massillon sandstone and Sierra White granite. One essential fact that comes to light during the analysis is that the effective shear moduli of the gas- and liquid-saturated regions must differ, otherwise it is impossible to account for the laboratory torsional wave data. Furthermore, the drainage analysis appears to give improved qualitative and quantitative agreement with the data for both of the materials considered. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1861712]

PACS numbers: 43.20.Jr, 43.35.Cg, 62.30.+d, 43.20.Bi [ANN]

Pages: 1785–1795

I. INTRODUCTION

The most successful and also most common method of remote sensing used in oil and gas exploration is reflection seismology.¹ The goals motivating use of such acoustic/seismic methods are first to delineate geologic traps for fluids and then to determine (to whatever extent possible) the nature and spatial distribution of the fluids in those traps. These problems are challenging from viewpoints of both engineering practice and fundamental science of acoustics. The present work addresses one unresolved fundamental issue concerning effects that nonuniform spatial distributions of pore fluids have on sound speeds in laboratory ultrasound experiments.

The classic work of Pochhammer² and Chree³ gave exact solutions for wave propagation in elastic rods. When the rod is instead a porous cylinder with fluid-filled pores, the equations of linear elasticity do not describe all possible motions of the fluid/porous-solid mixture. Biot's theory of fluid-saturated porous media^{4,5} provides a continuum theory, permitting the fluid and solid components to move independently and accounts approximately for the attenuation of waves due to viscous friction. Gardner^{6,7} used Biot's theory^{4,5} to study long-wavelength extensional waves in circular cylinders. Gardner considered only the low-frequency regime where the second bulk compressional mode predicted

by Biot's theory is diffusive in character. Gardner also limited consideration to the case of open-pore surface boundary conditions.

The present work is based in part on another paper by Berryman,⁸ in which both open- and closed-pore surface boundary conditions for the fluid-saturated porous cylinder were studied. Here we consider only the open-pore surface, but we allow patchy (nonuniform) saturation^{9–16} inside the cylinder. In particular, it is quite common to study partial saturation in the laboratory under drainage or drying conditions wherein an initially fully saturated porous cylinder is allowed to dry (from the outside in) while the experimenter continues to acquire data on the cylinder's modes of oscillation. We want to model this behavior explicitly, as we anticipate the results to differ significantly from those for homogeneous partial saturation. The simplest model of patchy saturation due to drainage from laboratory samples is surely concentric cylinders with a fully dry (air saturated) outer cylindrical shell enclosing a fully liquid-saturated inner cylinder.

No doubt a more realistic model would involve many shells with various degrees of partial saturation between the dry outer shell and the saturated inner cylinder at the core, but such complications will not be treated here. We find that studies of the two-layer case have all the important physical complications expected in this problem, while still having enough simplicity that some of the analysis can be done semi-analytically—thereby providing some of the sought after insight into the problem.

^{a)}Electronic mail: berryman1@llnl.gov

^{b)}Electronic mail: spride@lbl.gov

One somewhat unusual aspect of the presentation is that we allow the possibility that the shear modulus of a liquid-filled porous medium can be larger than that of the same medium filled with gas.¹⁷⁻¹⁹ Although this concept disagrees with Gassmann's quasi-static result for isotropic materials, it is now well understood that fluid dependence of shear is possible for heterogeneous, unrelaxed, and/or high-frequency experiments. In fact the laboratory data we are attempting to model and ultimately explain demand that this possibility be considered.

We present the equations of poroelasticity, and then show the forms of the equations needed for cylindrical geometry. Appropriate boundary conditions for our problem are discussed. Equations are subsequently formulated to determine both the extensional and torsional modes of concentric poroelastic cylinders under conditions of partial saturation. Solutions of these equations are computed and discussed here for torsional waves, while the harder problem of extensional waves will be treated fully in a second installment to be presented at a later date.

II. EQUATIONS OF POROELASTICITY

For long-wavelength disturbances (i.e., wavelength $\lambda \gg h$, where h is a typical pore size) propagating through a porous medium, we define average values of the (local) displacements in the solid and also in the saturating fluid. The average displacement vector for the solid frame is \mathbf{u} while that for the pore fluid is \mathbf{u}_f . The average displacement of the fluid relative to the frame is $\mathbf{w} = \phi(\mathbf{u}_f - \mathbf{u})$, where ϕ is porosity. For small strains, the frame dilatation is

$$e = e_x + e_y + e_z = \nabla \cdot \mathbf{u}, \quad (1)$$

where e_x, e_y, e_z are the Cartesian strain components. Similarly, the average fluid dilatation is

$$e_f = \nabla \cdot \mathbf{u}_f \quad (2)$$

(e_f also includes flow terms as well as dilatation) and the increment of fluid content is defined by

$$\zeta = -\nabla \cdot \mathbf{w} = \phi(e - e_f). \quad (3)$$

With these definitions, Biot²⁰ obtains the stress-strain relations in the form

$$\delta\tau_{xx} = He - 2\mu(e_y + e_z) - C\zeta, \quad (4)$$

and similarly (with permutations) for the other compressional components $\delta\tau_{yy}, \delta\tau_{zz}$, while

$$\delta\tau_{zx} = \mu \left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x} \right), \quad (5)$$

and again for $\delta\tau_{yz}, \delta\tau_{xy}$ for the other shear components. Finally, for the fluid pressure,

$$\delta p_f = M\zeta - Ce. \quad (6)$$

The $\delta\tau_{ij}$ (for $i, j = x, y, z$) are deviations from equilibrium of average Cartesian stresses in the saturated porous material and δp_f is similarly the isotropic pressure deviation in the pore fluid.

With time dependence of the form $\exp(-i\omega t)$, the coupled wave equations that incorporate (4)–(6) take the form

$$\begin{aligned} \omega^2(\rho\mathbf{u} + \rho_f\mathbf{w}) &= C\nabla\zeta - (H - \mu)\nabla e - \mu\nabla^2\mathbf{u}, \\ \omega^2(\rho_f\mathbf{u} + q\mathbf{w}) &= M\nabla\zeta - C\nabla e, \end{aligned} \quad (7)$$

where $\rho = \phi\rho_f + (1 - \phi)\rho_m$ is the bulk-density of the material and $q = \rho_f[\alpha/\phi + iF(\xi)\eta/\kappa\omega]$ is the effective density of the fluid in relative motion. The kinematic viscosity of the liquid is η , the permeability of the porous frame is κ ; the dynamic viscosity factor is given, for our choice of sign for the frequency dependence, by $F(\xi) = \frac{1}{4}\{\xi T(\xi)/[1 + 2T(\xi)/i\xi]\}$, where $T(\xi) = [\text{ber}'(\xi) - i\text{bei}'(\xi)]/[\text{ber}'(\xi) - i\text{bei}'(\xi)]$ and $\xi = (\omega h^2/\eta)^{1/2}$. The functions $\text{ber}(\xi)$ and $\text{bei}(\xi)$ are the real and imaginary parts of the Kelvin function. The dynamic parameter h is a characteristic length generally associated with and comparable in magnitude to the steady-flow hydraulic radius. The tortuosity $\alpha \geq 1$ (a pure number related to the average grain shapes and/or the pore topology) has been measured by Johnson *et al.*²¹ and has also been estimated theoretically by Berryman.²² It is also related to the electrical formation factor F by $\alpha = \phi F$.

The poroelastic moduli H , C , and M are given²³⁻²⁷ by

$$H = K + \frac{4}{3}\mu + (K_m - K)^2/(D - K), \quad (8)$$

$$C = K_m(K_m - K)/(D - K), \quad (9)$$

and

$$M = K_m^2/(D - K), \quad (10)$$

where

$$D = K_m[1 + \phi(K_m/K_f - 1)] \quad (11)$$

with K_m being the solid grain bulk modulus, K_f being the pore fluid bulk modulus, and ϕ being the porosity. Equations (8)–(11) are correct as long as the porous material can be considered homogeneous on the microscopic scale as well as the macroscopic scale, which we assume to hold for the present application.

To decouple the wave equations (7) into Helmholtz equations for the three modes of propagation for isotropic systems [fast longitudinal, slow longitudinal, and transverse (or shear)], we note that the displacements \mathbf{u} and \mathbf{w} can be decomposed as

$$\mathbf{u} = \nabla Y + \nabla \times \boldsymbol{\beta}, \quad \mathbf{w} = \nabla \psi + \nabla \times \boldsymbol{\chi}, \quad (12)$$

where Y, ψ are scalar potentials and $\boldsymbol{\beta}, \boldsymbol{\chi}$ are vector potentials. Substituting (12) into (7), we find (7) is satisfied if the following two pairs of equations hold:

$$(\nabla^2 + k_s^2)\boldsymbol{\beta} = 0, \quad \boldsymbol{\chi} = -\rho_f\boldsymbol{\beta}/q \quad (13)$$

and

$$(\nabla^2 + k_{\pm}^2)A_{\pm} = 0, \quad (14)$$

where k_s is the shear (or transverse) wave number, k_+ is the fast-longitudinal wave number, and k_- is the slow-longitudinal wave number. These wave numbers are defined by

$$k_s^2 = \omega^2(\rho - \rho_f^2/q)/\mu \quad (15)$$

and

$$k_{\pm}^2 = \frac{1}{2}\{b + f \mp [(b-f)^2 + 4cd]^{1/2}\}, \quad (16)$$

$$b = \omega^2(\rho M - \rho_f C)/\Delta, \quad c = \omega^2(\rho_f M - qC)/\Delta, \\ d = \omega^2(\rho_f H - \rho C)/\Delta, \quad f = \omega^2(qH - \rho_f C)/\Delta, \quad (17)$$

with $\Delta = MH - C^2$. This linear combination of scalar potentials has been chosen to be $A_{\pm} = \Gamma_{\pm} Y + \psi$, where

$$\Gamma_{\pm} = d/(k_{\pm}^2 - b) = (k_{\pm}^2 - f)/c. \quad (18)$$

With the identification (18) of the coefficients Γ_{\pm} , the mode decoupling is complete.

Equations (13) and (14) are valid for any choice of coordinate system. They may therefore be applied to boundary value problems with arbitrary symmetry. Biot's theory can therefore be applied to porous elastic cylinders in the next section.

III. EQUATIONS FOR POROUS CYLINDER

To work most easily in cylindrical geometry, we rewrite the stress-strain relations (4)–(6) in cylindrical coordinates. If z is the coordinate along the cylinder axis while r and θ are the radial and azimuthal coordinates, it is not difficult to show that

$$\delta\tau_{rr} = He - 2\mu(e_{\theta} + e_z) - C\zeta, \quad (19)$$

$$\delta\tau_{r\theta} = \mu\left(\frac{\partial u_{\theta}}{\partial r} - \frac{u_{\theta}}{r} + \frac{1}{r}\frac{\partial u_r}{\partial \theta}\right), \quad (20)$$

$$\delta\tau_{rz} = \mu\left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r}\right), \quad (21)$$

and (6) for δp_f remains unchanged. The stress increments $\delta\tau_{zz}$, $\delta\tau_{\theta\theta}$, and $\delta\tau_{\theta z}$ are not of direct interest in the present application. The dilatations are given by

$$e = e_r + e_{\theta} + e_z, \quad (22)$$

where

$$e_r = \frac{\partial u_r}{\partial r}, \quad e_{\theta} = \frac{u_r}{r} + \frac{1}{r}\frac{\partial u_{\theta}}{\partial \theta}, \quad e_z = \frac{\partial u_z}{\partial z}. \quad (23)$$

We redefine the vector potential β in terms of two scalar potentials according to

$$\beta = \hat{z}\beta_1 + \nabla \times (\hat{z}\beta_2), \quad (24)$$

where \hat{z} is a unit vector in the z direction and both β_i 's satisfy

$$(\nabla^2 + k_s^2)\beta_i = 0 \quad \text{for } i=1,2. \quad (25)$$

For the problem of interest here, there are two distinct regions: The first region is a circular cylinder centered at the origin, within which solutions of (14) and (25) must be finite at the origin. Solutions of these Helmholtz equations take the form

$$A_{\pm} = \alpha_{\pm} J_0(j_{\pm}) \exp[i(k_z z - \omega t)], \quad (26)$$

$$\beta_1 = \gamma_s J_0(j_s) \exp[i(k_z z - \omega t)], \quad (27)$$

$$\beta_2 = (\alpha_s / ik_z) J_0(j_s) \exp[i(k_z z - \omega t)], \quad (28)$$

where

$$j_{\pm} = k_{\pm} r, \quad j_s = k_s r, \quad (29)$$

and

$$k_{\pm r}^2 = k_{\pm}^2 - k_z^2, \quad k_{sr}^2 = k_s^2 - k_z^2. \quad (30)$$

J_0 is the zeroth order Bessel function of the first kind. The coefficients α_{\pm} , α_s , γ_s , are constants to be determined from the boundary conditions. (Note that the coefficients α_{\pm} and α_s should not be confused with the tortuosity symbol α introduced earlier.)

The second region is a cylindrical shell surrounding the first region. In this region, the factors k_{\pm} and k_s can take different values from those in the central region, indicated by k_{\pm}^* and k_s^* (where $*$ means air-filled, and does *not* ever mean complex conjugate in this paper). Furthermore, two linearly independent solutions of the equations are allowed, i.e., both J_0 and Y_0 (the Bessel function of the second kind, sometimes known as the Neumann function). In the outer shell, we therefore have four coefficients apiece for J_0 and Y_0 , all of which must also be determined by the boundary conditions.

First note that

$$Y = (A_+ - A_-)/(\Gamma_+ - \Gamma_-), \\ \psi = (A_+ \Gamma_- - A_- \Gamma_+)/(\Gamma_- - \Gamma_+) \quad (31)$$

from the definitions of A_{\pm} . Then, substituting (26)–(28) and (31) into (12), and the result into (6) and (19)–(21), we learn how the stress distribution in the cylinder is related to the unknown potential amplitudes α_{\pm} , α_s , and γ_s :

$$\delta\tau_{r\theta} = m_{11}\gamma_s \equiv -\mu k_{sr}^2 J_2(j_s) \gamma_s, \quad (32)$$

$$\delta\tau_{rr} = a_{11}\alpha_+ + a_{12}\alpha_- + a_{13}\alpha_s, \quad (33)$$

$$-\delta p_f = a_{21}\alpha_+ + a_{22}\alpha_- + a_{23}\alpha_s, \quad (34)$$

$$\delta\tau_{rz} = a_{31}\alpha_+ + a_{32}\alpha_- + a_{33}\alpha_s. \quad (35)$$

Here the a_{ij} coefficients are defined in terms of known quantities—while using the definition $\Xi \equiv 1/(\Gamma_+ - \Gamma_-)$ —by

$$a_{11} = \Xi[(C\Gamma_- - H)k_+^2 + 2\mu k_z^2] J_0(j_+) \\ + 2\Xi \mu k_{+r} J_1(j_+)/r, \quad (36)$$

$$a_{12} = -2\Xi \mu k_{-r} J_1(j_-)/j_- \\ + \Xi[(H - C\Gamma_+)k_-^2 - 2\mu k_z^2] J_0(j_-), \quad (37)$$

$$a_{13} = 2\mu k_{sr}^2 J_2(j_s), \quad (38)$$

$$a_{21} = \Xi(M\Gamma_- - C)k_+^2 J_0(j_+), \quad (39)$$

$$a_{22} = \Xi(C - M\Gamma_+)k_-^2 J_0(j_-), \quad (40)$$

$$a_{31} = -2i\Xi \mu k_z k_{+r} J_1(j_+), \quad (41)$$

$$a_{32} = 2i\Xi \mu k_z k_{-r} J_1(j_-), \quad (42)$$

$$a_{33} = -\mu(k_s^2 - 2k_z^2)k_{sr} J_1(j_s)/ik_z, \quad (43)$$

and $a_{23}=0$. There is an implicit factor of $\exp[i(k_z z - \omega t)]$ on the right-hand side of (32)–(35).

Berryman⁸ has shown that a_{11} , a_{13} , a_{31} , and a_{33} reduce in the limit $\phi \rightarrow 0$ to the corresponding results for isotropic elastic cylinders by Pochhammer,² Chree,^{3,28} Love,²⁹ and Bancroft,³⁰ as they should.

IV. BOUNDARY CONDITIONS

Appropriate boundary conditions for use with Biot's equations have been considered by Deresiewicz and Skalak,³¹ Berryman and Thigpen,³² and Pride and Haartsen,³³ so we make use of these results here.

At the external surface $r=R_2$ where the outer porous material contacts the surrounding air, it is most appropriate for the drainage experiments to be analyzed while using free surface conditions

$$-\delta p_f = 0, \quad \delta \tau_{rr} = 0, \quad \delta \tau_{r\theta} = 0, \quad \delta \tau_{rz} = 0 \quad (44)$$

for the deviations from static equilibrium. (If instead the cylinder is sealed on $r=R_2$, then the first of these needs to be replaced by $w_r=0$. While of theoretical interest, this condition is not one appropriate for the conditions of the actual laboratory experiments we will analyze.)

The internal interface condition at $r=R_1$ needs more precise definition. We assume that all the meniscii that are separating the inner fluid (liquid) from the outer fluid (gas) are contained within a thin layer (shell) of thickness δh (a few grain sizes in width) straddling the surface $r=R_1$. The porous-continuum boundary conditions used at the $r=R_1$ interface are to be understood in the limit where $\delta h/R_1 \rightarrow 0$.

All fluid that enters this interface layer goes into stretching the meniscii since as Pride and Flekkoy³⁴ have shown, it is reasonable to assume that the contact lines of the meniscii remain pinned under linear seismic stressing (strains $\leq 10^{-6}$). The locally incompressible flow conserves fluid volume so that the rate at which the inner fluid enters the interface layer is equal to the rate at which the outer fluid leaves the layer. With the assumed $e^{-i\omega t}$ time dependence, this becomes the fluid displacement condition

$$w_r(r=R_1 + \delta h/2) = w_r^*(r=R_1 - \delta h/2). \quad (45)$$

Another boundary condition defines the incremental change in the fluid-pressure difference across the interface layer as being the surface tension times the average incremental change in the curvature of the meniscii. In terms of porous-continuum variables, Nagy and Blaho³⁵ and Nagy and Nayfeh³⁶ have defined this interface condition as

$$\delta p_f - \delta p_f^* = W w_r, \quad (46)$$

where W is called the membrane stiffness. Nagy and Nayfeh³⁶ have shown that $W = \sigma/\kappa$ (where σ is surface tensions and κ is permeability) for pore models in which the pores are straight cylinders. Pride *et al.*³⁷ have shown that when the dimensionless inequality $\sigma R_2^2/(\kappa K R_1) < 1$ is satisfied, finite stiffness of the meniscii may be neglected so that

$$\delta p_f = \delta p_f^* \quad (47)$$

everywhere along the interface $r=R_1$ (recall that R_2 is the sample radius). With $\kappa = 100$ mDarcy, $K = 10$ GPa, and σ

$= 10^{-2}$ Pa m (typical values for a sandstone and an air/water interface), we find that R_2^2/R_1 should be less than 10 cm in order for Eq. (47) to apply. Since most laboratory samples are smaller than 10 cm in radius, Eq. (47) is therefore valid at all saturation levels for the types of materials of interest here, so no further reference to the surface tension (or membrane stiffness) needs to be made.

We also have the standard porous-continuum results that

$$\tau_{rr} = \tau_{rr}^*, \quad \tau_{r\theta} = \tau_{r\theta}^*, \quad \tau_{rz} = \tau_{rz}^*, \quad (48)$$

which express momentum conservation at the interface, and the time derivatives of the displacement components satisfy

$$\dot{u}_r = \dot{u}_r^*, \quad \dot{u}_\theta = \dot{u}_\theta^*, \quad \dot{u}_z = \dot{u}_z^*, \quad (49)$$

which express the fact that the grains remain in (welded) contact at the interface.

To apply the boundary conditions (45) and (47), we need in addition to (34) the result

$$w_r = a_{41}\alpha_+ + a_{42}\alpha_- + a_{43}\alpha_s, \quad (50)$$

where [again using $\Xi \equiv 1/(\Gamma_+ - \Gamma_-)$]

$$a_{41} = \Xi k_{+r} J_1(j_+) \Gamma_-, \quad (51)$$

$$a_{42} = -\Xi k_{-r} J_1(j_-) \Gamma_+, \quad (52)$$

$$a_{43} = k_{sr} J_1(j_s) \rho_f / q. \quad (53)$$

The remaining stress conditions (48) are determined by (33) and (35).

To apply the boundary conditions (49), we need the explicit expressions for the displacement which follow from (12). The results are of the form

$$u_r = a_{51}\alpha_+ + a_{52}\alpha_- + a_{53}\alpha_s, \quad (54)$$

where

$$(\Gamma_+ - \Gamma_-) a_{51} = -k_{+r} J_1(j_+), \quad (55)$$

$$(\Gamma_+ - \Gamma_-) a_{52} = k_{-r} J_1(j_-), \quad (56)$$

$$a_{53} = k_{sr} J_1(j_s), \quad (57)$$

and

$$u_z = a_{61}\alpha_+ + a_{62}\alpha_- + a_{63}\alpha_s, \quad (58)$$

where $a_{61} = a_{62} = 0$ and

$$a_{63} = k_{sr}^2 J_0(j_s) / i k_z. \quad (59)$$

Both (54) and (58) are needed for extensional waves, while the remaining component,

$$u_\theta = m_{21} \gamma_s \equiv k_{sr} J_1(j_s) \gamma_s, \quad (60)$$

is needed only for torsional waves. As before, there is an implicit factor of $\exp i(k_z z - \omega t)$ on the right-hand side of (51)–(53), (55)–(57), and (59).

It follows from (32)–(35), (50), and (60) that γ_s (for the inner cylinder) and the corresponding coefficients for the cylindrical shell are all completely independent of the other mode coefficients and, therefore, relevant to the study of torsional waves, but not for extensional waves. Pertinent equations for the torsional wave dispersion relation are continuity

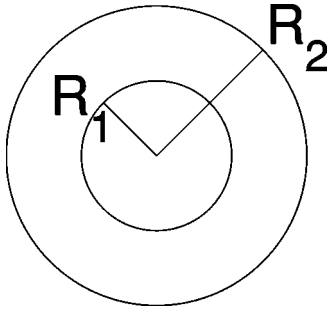


FIG. 1. Cross-section of a circular cylinder, where $R_1 = S^{1/2}R_2$ is determined by the liquid saturation level S .

of the angular displacement, u_θ , and stress, $\tau_{r\theta}$, at the internal interface, and vanishing of the stress, $\tau_{r\theta}$, at the external surface.

The final set of equations for the extensional wave dispersion relation involves nine equations with nine unknowns. The nine unknowns are α_+ , α_- , α_s (coefficients of J_0 in the central cylinder), plus three α^* 's (coefficients of J_0) and three η^* 's (coefficients of Y_0) for region of the cylindrical shell. The nine equations are the continuity of radial and one tangential stress as well as radial and one tangential displacement at the interfacial boundary (totaling four conditions), continuity of fluid pressure and normal fluid increments across the same boundary (two conditions), and finally the vanishing of the external fluid pressure, and radial and one tangential stress at the free surface (three conditions). The extensional wave dispersion relation is then determined as in Berryman⁸ by those conditions on the wave number k_z that result in vanishing of the determinant of the coefficients of this 9×9 complex matrix.

V. ELEMENTARY TORSIONAL MODES

The torsional mode of cylinder oscillation (which is trivial for a simple cylinder, porous or not) is determined here by a 3×3 system, of which eight elements are in general nonzero. This system is therefore similar in size and difficulty to the cases studied earlier by Berryman⁸ for extensional waves in a simple fully saturated poroelastic cylinder. On the other hand, for extensional waves, the matrix determining the extensional wave dispersion relation for patchy saturation has 81 elements, of which 69 will in general be nonzero. This problem requires sufficiently different treatment from that for the torsional case that we set it aside to be studied fully in a separate presentation (Part II).

We assume that the cylinder has liquid saturation level $S = (R_1/R_2)^2$, where R_2 is the radius of the cylinder and $r = R_1$ is the location of the liquid–gas interface (see Sec. IV and Fig. 1). The dispersion relation for torsional waves is then given by

$$\begin{vmatrix} m_{11}^*(R_2) & n_{11}^*(R_2) & 0 \\ -m_{11}^*(R_1) & -n_{11}^*(R_1) & m_{11}(R_1) \\ -m_{21}^*(R_1) & -n_{21}^*(R_1) & m_{21}(R_1) \end{vmatrix} = 0, \quad (61)$$

where m_{11} and m_{21} are given by (32) and (60). The coefficients m_{11}^* and m_{21}^* have the same functional forms as m_{11}

and m_{21} , but the constants are those for the shell, rather than the inner cylinder. Similarly, n_{11}^* and n_{21}^* are just the same as m_{11}^* and m_{21}^* except that J_0 and J_1 are replaced everywhere by Y_0 and Y_1 , respectively.

Analysis proceeds by noticing immediately that there could be two elementary solutions of (61), one with $m_{11}^*(R_2) = n_{11}^*(R_2) = 0$ (exterior condition) and another with $m_{11}(R_1) = m_{21}(R_1) = 0$ (interior condition). First, the interior condition is satisfied, for example, when $k_{sr} = 0$ or, equivalently, when $k_z^2 = k_s^2$. This corresponds to a torsional mode of propagation having wave speed and attenuation determined exactly by the bulk shear wave number in the saturated interior region, but this interior region is not moving since $k_{sr} = 0$ also implies that $u_\theta = 0$ from (60). Thus, the interior condition results in the drained outer shell twisting back and forth around a stationary inner liquid-saturated cylinder. Second, the exterior condition is similarly satisfied when $k_{sr}^* = 0$ or, equivalently, when $k_z^2 = (k_s^*)^2$. This condition looks at first glance as if it might be spurious because $k_{sr}^* = 0$ suggests that u_θ at the exterior boundary might vanish identically, and then this would correspond to a trivial solution of the equations. However, looking closer, this is not the case, because at the external boundary

$$u_\theta = k_{sr}^* [J_1(j_s^*) \gamma_s^* + Y_1(j_s^*) \epsilon_s^*], \quad (62)$$

so as $k_{sr}^* \rightarrow 0$, the first term on the right-hand side of (62) does vanish, both because $k_{sr}^* \rightarrow 0$ and also because $J_1(j_s^*) \rightarrow 0$. But the second term does not vanish in this limit because $|Y_1(j_s^*)| \rightarrow 2/\pi k_{sr}^* R_2 \rightarrow \infty$ as $k_{sr}^* \rightarrow 0$, and the product gives the finite result: $2/\pi R_2$. So this condition is not spurious and corresponds to a torsional wave propagating with the speed and attenuation of the bulk shear wave number in the outer, drained shell material.

Can both of these elementary modes be excited? If we assume for the moment that Gassmann's equations²³ (also see Berryman³⁸) apply to the sample, then $\mu^* = \mu$, and the only changes in shear wave velocity in the two regions are those induced by the changes in mass. In this situation, the wave speed in the air/gas-saturated region will be faster than that in the water/liquid-saturated region, since liquid is more dense than gas. Thus, the real part of k_s^* is smaller than that of k_s , and while the condition $(k_{sr}^*)^2 = 0$ implies that the real part of k_{sr}^* is positive, the condition $k_z^2 = 0$ implies that the real part of $(k_{sr}^*)^2$ is negative. Therefore, assuming (as we generally do here) that the attenuation in the system is relatively small, the condition $k_z = k_s^*$ leads to a propagating wave, while $k_z = k_s$ leads to a strongly evanescent wave. Thus, for Gassmann conditions on shear, only one of these possible modes actually propagates.

Note that, if/when Gassmann's results do not apply to the system (say at ultrasonic frequencies^{17–19,39}), then the results of the preceding paragraph must necessarily be reconsidered. In particular, if the shear modulus changes rapidly with the introduction of liquid saturant, it is possible that the shear wave speed for a liquid-saturated porous material may be higher than that for the gas saturated case. In this situation, all the inequalities of the preceding paragraph would be reversed, and then the condition $k_z = k_s$ leads to a propagating wave, while $k_z = k_s^*$ leads to a strongly evanescent wave.

TABLE I. The zeroes $j_{2,n}$ of $J_2(z)$ as a function of the order n of appearance along the real axis.

Order n	$j_{2,n}$
0	0.000 000 000 000 00
1	5.135 622 301 840 68
2	8.417 244 140 399 87
3	11.619 841 172 149 06

Our conclusion then is that both modes can indeed be excited, but probably not simultaneously in the same system and/or in the same frequency band. In a highly dispersive porous system and with broadband acoustic signal input, it could happen that both modes are propagating simultaneously in time, but in distinct/disjoint frequency bands.

VI. HIGHER ORDER TORSIONAL MODES

For fully saturated porous cylinders, the factor that determines the torsional modes of propagation is $m_{11}(r)$ in (32). The critical factor here is the Bessel function $J_2(j_s)$ and, specifically, the whereabouts of its zeroes. One source of this information, to five figure accuracy, is the reference of Abramowitz and Stegun,⁴⁰ which provides not only the location of the zeroes $j_{2,n}$, but also the values of the corresponding derivatives $J_2'(j_{2,n})$. Having these derivatives is useful for improving the accuracy of the zeroes with a Newton–Raphson iterative method, based on $j_{2,n} = j_{2,n}^{old} - J_2(j_{2,n}^{old})/J_2'(j_{2,n}^{old})$. This approach gives a very rapid improvement to the values of the $j_{2,n}$'s within two to three iterations. The results to order $n=3$ are shown in Table I.

Having already understood the zeroth-order contributions to the dispersion relation (61) due to zeroes of k_{sr} and k_{sr}^* , we are now free to consider that neither of these factors vanishes for the higher-order modes. This assumption permits us to factor these wave numbers in or out of the determinant whenever it is convenient to do so. In particular, we note that the first two columns of (61) would have a common factor of $\mu^*(k_{sr}^*)^2$ (which could then be safely eliminated) if we first multiply the bottom row by a factor of $\mu^*k_{sr}^*$. Having made these simplifications, we find

$$\begin{vmatrix} J_2(k_{sr}^*R_2) & Y_2(k_{sr}^*R_2) & 0 \\ J_2(k_{sr}^*R_1) & Y_2(k_{sr}^*R_1) & \mu k_{sr} J_2(k_{sr}R_1) \\ J_1(k_{sr}^*R_1) & Y_1(k_{sr}^*R_1) & \mu^*k_{sr}^*J_1(k_{sr}R_1) \end{vmatrix} = 0, \quad (63)$$

after also eliminating a common factor of -1 from the top row, and $-k_{sr}$ from the third column.

Expanding the determinant along the third column, we have

$$0 = \mu^*k_{sr}^*J_1(k_{sr}R_1) \begin{vmatrix} J_2(k_{sr}^*R_2) & Y_2(k_{sr}^*R_2) \\ J_2(k_{sr}^*R_1) & Y_2(k_{sr}^*R_1) \end{vmatrix} - \mu k_{sr} J_2(k_{sr}R_1) \begin{vmatrix} J_2(k_{sr}^*R_2) & Y_2(k_{sr}^*R_2) \\ J_1(k_{sr}^*R_1) & Y_1(k_{sr}^*R_1) \end{vmatrix}. \quad (64)$$

Some elementary consequences of this equation are the following. (a) As $R_1 \rightarrow 0$ so there is no liquid left in the system,

$J_1(k_{sr}R_1)$ and $J_2(k_{sr}R_1) \rightarrow 0$ like R_1 , while $Y_1(k_{sr}^*R_1)$ and $Y_2(k_{sr}^*R_1) \rightarrow \infty$ like $1/R_1$. So the dispersion relation is always satisfied in the limit when $J_2(k_{sr}^*R_2) = 0$, which is exactly the condition for the fully dry cylinder as expected. (b) If $R_1 \rightarrow R_2$, then the first determinant vanishes identically. The second determinant does not vanish in general since it approaches the Wronskian $J_2Y_1 - J_1Y_2 = 2/\pi k_{sr}^*R_2$, so the condition becomes $k_{sr}J_2(k_{sr}R_1) = 0$, again as expected. (c) The special case of $k_{sr} \rightarrow 0$ does not affect these conclusions, as both $J_1(k_{sr}R_1)$ and $J_2(k_{sr}R_1) \rightarrow 0$ in this limit, as they should. (d) The only case that is missing from (64) is the one for $k_{sr}^* \rightarrow 0$. But this multiple zero of the original dispersion relation (61) was eliminated when we removed two factors of $(k_{sr}^*)^2$ from the first and second column in the first step of our simplification of the dispersion relation – a step which is always legitimate *except* when $k_{sr}^* = 0$.

We conclude that, with the one trivial exception just noted, these simplifications have kept the basic nature of the dispersion relation intact.

A. Lower frequency results

At lower frequencies in the range $f < 1$ kHz, we may typically expect^{18,41,42} that Gassmann's results hold for the poroelastic medium, where $\mu^* \equiv \mu$. Also, to a very good approximation $k_s^* \simeq k_s$, where the only deviations from equality are those due to the differences in the densities of liquid and gas constituents. So deviations from this approximation are most substantial when the porosity is high. From (63), we see that if the products μk_{sr} and $\mu^*k_{sr}^*$ are equal, then these factors can be removed from the third column of the determinant. Then, the resulting third column can be subtracted from the first column, and the result can be expanded along the first column to give

$$\frac{2J_2(k_{sr}R_2)}{\pi k_{sr}R_1} \simeq 0, \quad (65)$$

having again used the fact that $J_2(z)Y_1(z) - J_1(z)Y_2(z) = 2/\pi z$. So the important zeroes in this case are again those of J_2 , some of which are already displayed in Table I.

Ignoring the imaginary part of k , which is usually quite small in this limit, we have the analytical result that

$$v_z^{(n)} = \frac{v_s}{[1 - (j_{2,n}v_s/\omega R_2)^2]^{1/2}}. \quad (66)$$

Thus, at the higher frequencies, this velocity approaches that of the shear wave as expected. When the lower frequencies are approached, there is an obvious cutoff frequency, $f_c^n = j_{2,n}v_s/2\pi R_2$, below which these torsional modes do not propagate for $n \geq 1$. Since this low-frequency cutoff may often be inconsistent with the assumption under consideration here (i.e., that frequencies are low enough so Gassmann's equation is satisfied), we expect generally that very few of the higher-order modes can be excited in this limit. The main result is therefore that $v_z = v_z^0 = v_s$ is the velocity that will normally be observed in laboratory experiments in this frequency domain, with only very rare exceptions. Recent experimental findings of Wisse *et al.*,⁴³ although for a somewhat different physical problem, nevertheless seem to be consistent with these results.

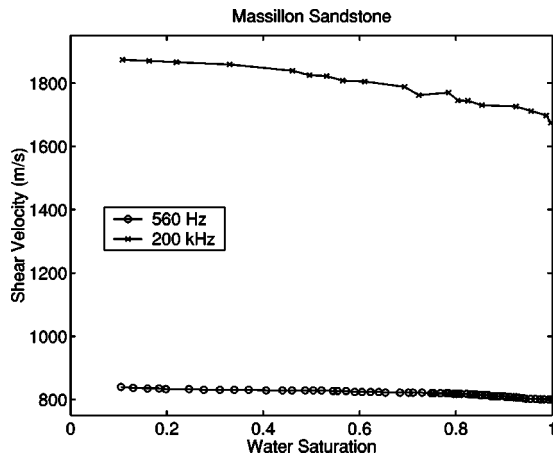


FIG. 2. Shear wave velocities as a function of water saturation for drainage experiments by Murphy.^{44,45}

We will not consider this rather special case any further in this paper.

B. Higher-frequency results

The more interesting case is that for higher frequencies, in which case it is now understood^{18,41,42} that the simple Biot–Gassmann theory is actually inadequate because there can be dependence of μ on liquid saturant properties at higher frequencies, in the range $f > 1$ kHz. The precise frequency at which this becomes important is material dependent and, therefore, the value 1 kHz is merely a convenient rule of thumb, but it is generally observed that for ultrasonic frequencies $f > 20$ kHz some deviations from Biot–Gassmann predictions are normally found. So it is in this regime that the distinctions between air- and water-filled pores become important for the torsional motion of a cylinder.

1. Case: $\mu^* = \mu$

Even if there is no difference between μ^* and μ , there can still be significant differences between k_s^* and k_s due to the differences in the fluid viscosities and densities of liquid and gas constituents. So we discuss this case now. For Massillon sandstone, Murphy^{44,45} measured extensional and shear wave velocities at $f = 560$ Hz and at $f = 200$ kHz over a range of partial saturations produced using the drainage method (Fig. 2). Relevant properties of this sandstone are listed in Table II. Data for the lower-frequency (560 Hz) experiment appear to satisfy the condition $\mu^* = \mu$, as has often been observed about this data set.

TABLE II. Properties of Massillon sandstone used in Murphy’s experiments¹¹ and Spirit River sandstone in Knight and Nolen-Hoeksema’s experiments.¹³

Property	Massillon	Spirit River
Porosity (%)	23.0	5.2
Permeability (mD)	7.37×10^2	1.0×10^{-3}
Grain size (μm)	150–200	125–150

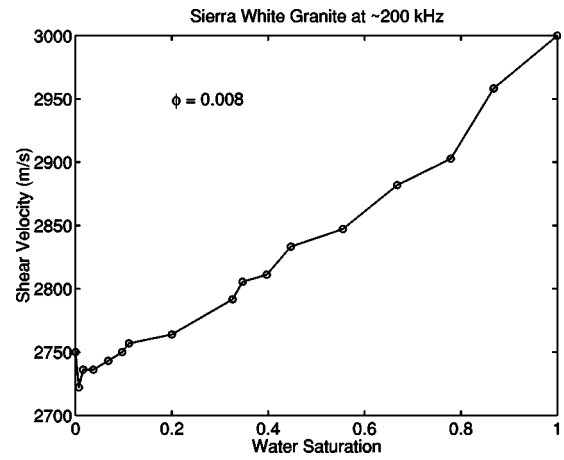


FIG. 3. Shear wave velocities as a function of water saturation for drainage experiments by Murphy⁴⁴ in Sierra White granite.

2. Case: $\mu^* \leq \mu$

The presence of liquid in the pores may alter the mechanical behavior of rocks under shear deformations in at least two quite distinct ways: (a) It is often observed that a very small amount of some liquids can cause chemical interactions that tend to soften the binding material present among the grains of such a system. When this happens, the shear modulus is usually observed to decrease. (See, for example, Fig. 3 for Sierra White granite.⁴⁴) So this situation implies that $\mu^* \geq \mu$, contrary to both Gassmann’s results and the stiffening effect that might be expected due to having liquid in the pores. Although this situation is well known in practice, we will ignore it in our modeling efforts. Our justification for this will be that the medium we are calling “dry” should in fact be termed “drained” (but with the bulk of the liquid actually removed from the pores) in the sense that it has been wetted previously and therefore has these chemical softening effects already factored into the modulus μ^* . In any case, our goal here is not so much to fit data for specific rocks, but rather to understand general trends. (b) The other situation that can also occur in practice—particularly at higher frequencies—is that the liquid saturating the porous material can have a non-negligible mechanical effect^{17–19,42,46} that tends to strengthen the medium under shear loading so that $\mu \geq \mu^*$. If this strengthening effect is great enough [and there are experimental results (see Fig. 4) that confirm this does happen in practice¹⁸], then it is possible the density effect is more than counterbalanced by the enhanced shear modulus effect with the result that the speed of shear wave propagation in the liquid-saturated medium is greater than that in the air-saturated case. Depending on details of the liquid distribution in the pores, either of these cases can be included in the analysis.

For Massillon sandstone, Murphy^{44,45} also measured extensional and shear wave velocities at $f = 200$ kHz over a range of partial saturations produced using the drainage method. Relevant properties of this sandstone were listed before in Table II.

VII. SOLVING THE DISPERSION RELATION

We will now show results for two cases (both from Murphy’s thesis⁴⁴): first the Massillon sandstone (at 560 Hz) and

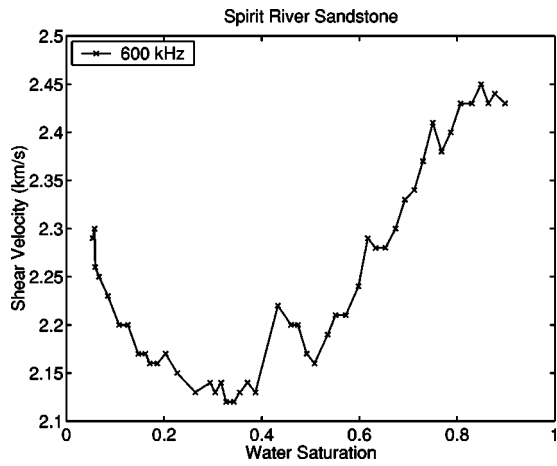


FIG. 4. Shear wave velocities as a function of water saturation for drainage experiments of Knight and Nolen-Hoeksema¹³ in Spirit River sandstone.

then the Sierra White granite (at about 200 kHz). We might expect based just on the experimental frequencies that the sandstone behavior would be close to that predicted by Gassmann, while that of the granite may differ from Gassmann.

An important observation concerning how to proceed with the analysis follows from the fact that we are seeking a curve in the complex plane, points along the curve depending on the level of saturation S . We know (at least in principle) the locations of the end points of this curve since they are exactly the points for full liquid saturation and full gas saturation. If we assume that the attenuation is relatively small so the wave numbers k_s and k_s^* have small imaginary parts, then to a reasonable approximation it must be the case that the curve of interest lies close to the real axis in the complex k_z^2 plane. If the imaginary parts exactly vanish, the curve reduces to a straight line on the real axis in this plane. These observations suggest that it might be helpful to trace rays in the complex plane radiating out from the origin, and in particular a ray (i.e., a straight line) passing through the origin and also through the point corresponding to whichever point, k_s^2 or $(k_s^*)^2$, happens to lie closest to the origin should provide a good starting point for the analysis. Another alternative is to consider the straight line that connects these two points directly, even though it would not in general also be a ray through the origin (unless there is no attenuation). Both of these alternatives have been tried.

The first alternative, considering a complex ray through the origin and then passing through the closest point k_s^2 or $(k_s^*)^2$, has the very important characteristic that the values of the dispersion function become purely imaginary in the shadow (i.e., in the contiguous region leading away from the origin) of the starting point of the curve. This fact provides a great simplification because we need the dispersion function to vanish identically—both in real and imaginary parts, and this shadow region has the nice characteristic that the real part is automatically zero. So the only remaining issue is to check where the imaginary part vanishes. This procedure is much easier to implement and to understand intuitively than trying to find the complex zeroes using something like a Newton method, which could also be implemented for this problem.

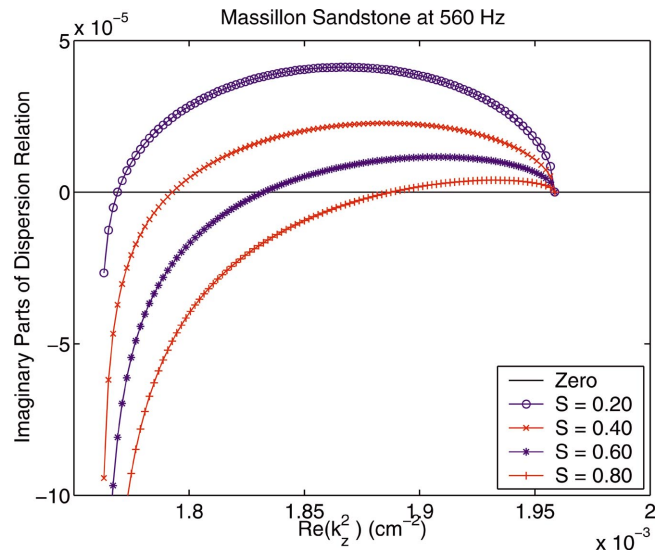


FIG. 5. Showing how the imaginary parts of the dispersion relation for Massillon sandstone change in the complex k_z^2 plane as k_z varies from k_{sa} to k_{sw} . The real part of the dispersion relation is either zero or very close to zero along this line and therefore the desired points are those where the imaginary part crosses the zero line.

The second alternative is not as rigorous as the first, but for the case of small attenuation gives very similar results and is especially easy to implement. In this case we need only consider the line connecting the two points k_s^2 and $(k_s^*)^2$ in the complex plane. It turns out that in the two cases considered here, the real part of the dispersion function is again either zero or very small, so that it makes sense to treat this approach as an approximation to the first one in that we need only seek the points where the imaginary part vanishes. This procedure is very intuitive and examples are shown in Figs. 5–8.

A. Massillon sandstone

For Massillon, we have the Gassmann-like situation in which the shear wave speed for the drained case is smaller

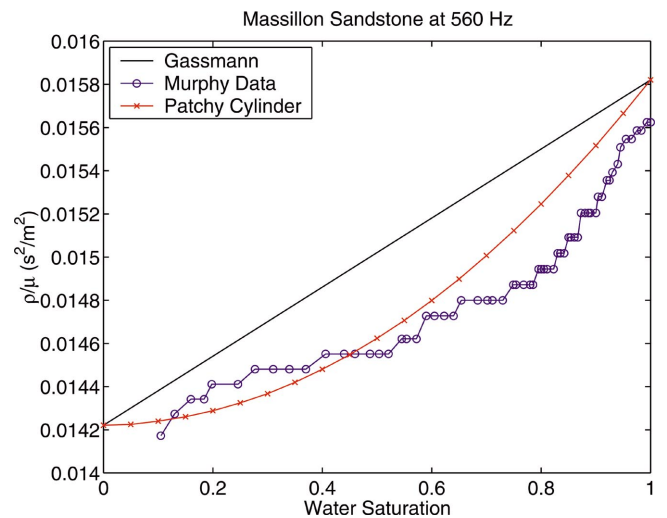


FIG. 6. Comparison between the points that solve the dispersion relation for the patchy cylinder, plotted as $\rho/\mu = 1/v_s^2$ versus water saturation S , for Massillon sandstone at 560 Hz. Data are from Murphy.^{44,45} The Gassmann curve is computed assuming that the shear modulus μ is constant and that the only quantity changing is therefore the density ρ .

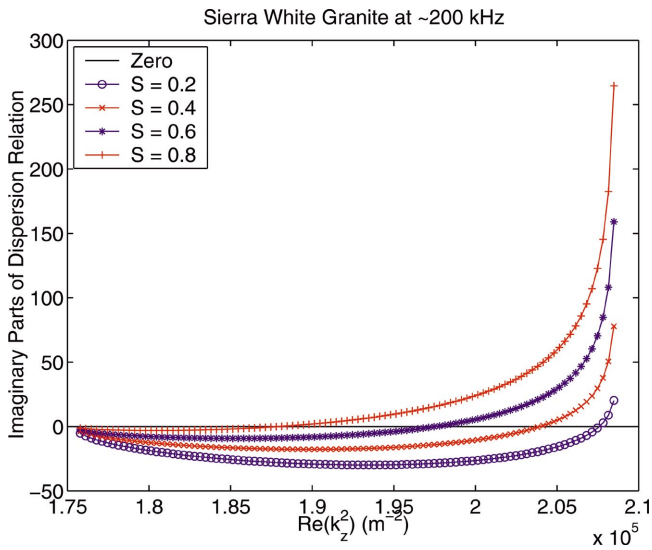


FIG. 7. Showing how the imaginary parts of the dispersion relation for Sierra white granite change in the complex k_z^2 plane as k_z varies from k_{sd} to k_{sw} . The real part of the dispersion relation is either zero or very close to zero along this line and therefore the desired points are those where the imaginary part crosses the zero line.

than that for the fully saturated case and therefore $\text{Re}(k_s^*) < \text{Re}(k_s)$. Figure 5 shows how the imaginary parts of the dispersion function change in this case as the real part of k_z^2 varies from $\text{Re}[(k_s^*)^2]$ to $\text{Re}(k_s^2)$ (i.e., from air saturated to water saturated). Figure 5 shows four of these curves ($S = 0.2$ to 0.8). Figure 6 was generated by completing the procedure for 19 equally spaced points in saturation S . Figure 6 shows, furthermore, that the curve obtained actually fits the data for Massillon better than does Gassmann (i.e., represented by the straight line between the end points). This result was a bit of a surprise as virtually everyone (including

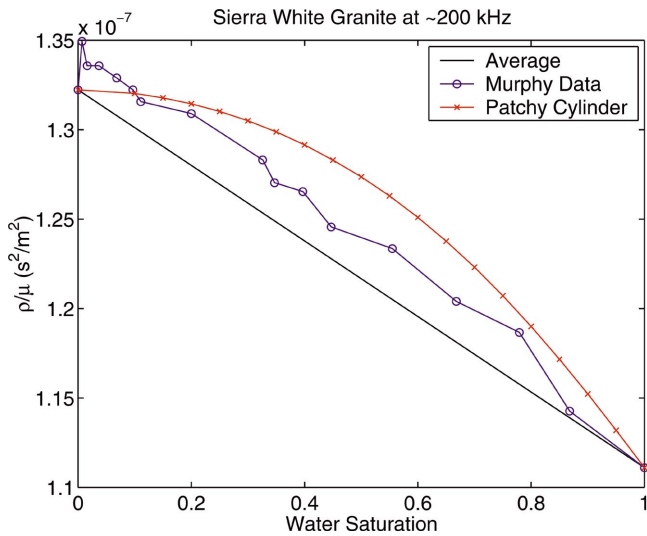


FIG. 8. Comparison between the points that solve the dispersion relation for the patchy cylinder, plotted as $\rho/\mu = 1/v_s^2$ versus water saturation S , for Sierra White granite at 200 kHz. Data are from Murphy.^{44,45} For Sierra White, Gassmann's equation clearly does not apply since the shear modulus μ must have increased with water saturation. Data and patchy calculation results are therefore compared to the saturation weighted average of $1/v_s^2$ in analogy to the Gassmann result.

the present authors) has often considered these data to be the best known proof of the accuracy of Gassmann's equations for partial saturation problems.

In fact, we could have chosen to fit the data better in the region close to full saturation. This choice would have moved the two theoretical curves down somewhat, but nevertheless this would not change our general conclusions that the data do not fall on a straight line, and that the trend in the data is more closely approximated by the patchy saturation curve. The slope of the Gassmann straight line depends only on the solid and fluid densities, and (except for a fixed scale factor) not on the shear modulus of the porous solid.

B. Sierra White granite

For Sierra White (data from Murphy⁴⁴), we have the non-Gassmann-like situation in which the shear wave speed for the drained case is larger than that for the fully saturated case and therefore $\text{Re}(k_s^*) > \text{Re}(k_s)$. Figure 7 shows how the imaginary parts of the dispersion function change in this case as the real part of k_z^2 varies from $\text{Re}(k_s^2)$ to $\text{Re}[(k_s^*)^2]$ (i.e., from water saturated to air saturated). Figure 7 shows four of these curves ($S = 0.2$ to 0.8). Figure 8 was generated by completing the procedure for 19 equally spaced points in saturation S . Figure 8 shows, furthermore, that both data and the curve obtained here differ substantially from the simple straightline average that might have been anticipated. Furthermore, the dispersion curve does in fact move in the right direction to agree with the data. What to expect in this situation was certainly not known by the authors at the outset since the common understanding of poroelasticity does not extend to this rather difficult set of partial saturation problems. But, it is gratifying to see this relatively simple theory clearly picks the right trends and agrees reasonably well with these data.

C. General behavior of the curves

Since the curves obtained in Figs. 6 and 8 are very well behaved, it seems appropriate to check for simple dependencies on the saturation parameter S . Both curves look like they might be linear in S with a small amplitude quadratic correction. This hypothesis is tested in Figs. 9 and 10. We find that the quadratic dependence is essentially exact to graphical accuracy for Massillon sandstone, and it is close but not exact for the Sierra White granite. This result is intriguing, but we have not yet tried to analyze this behavior and therefore will not attempt an explanation of it at this time.

VIII. SUMMARY

Biot fast- and slow-wave effects in layered materials have been studied previously by Pride *et al.*⁴⁷ and by many others found in their references. The present work is motivated by the desire to understand how fluids interacting with common poroelastic systems may result in observed wave speeds and viscous attenuation in partially saturated (and especially in patchy saturated) cylinders. These effects can then be observed in the speeds and attenuations of extensional and torsional waves. There are large quantities of such data al-

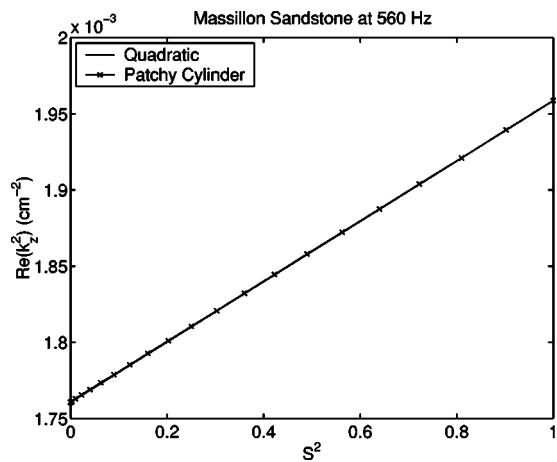


FIG. 9. Comparison between the points that solve the dispersion relation in the complex k_z^2 plane and a simple quadratic fit in water saturation S , for the Massillon sandstone measured by Murphy.^{44,45} The fit is nearly perfect for this case.

ready available, and one thrust of our future work will be to reanalyze these data in light of the methods developed here. We have concentrated on analysis of the shear and torsional wave speeds here, as this is clearly the first essential step in the overall analysis of these problems. The next step will be the more complicated solution of the extensional wave problem for these same systems.

ACKNOWLEDGMENTS

We thank Bill Murphy and Rosemarie Knight for providing access to their unpublished data files. We thank Brad Artman for helpful comments on the manuscript. The work of JGB was performed under the auspices of the U.S. Department of Energy under contract No. W-7405-ENG-48 and supported specifically by the Geosciences Research Program of the DOE Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences. The work of SRP was supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC03-76F00098.

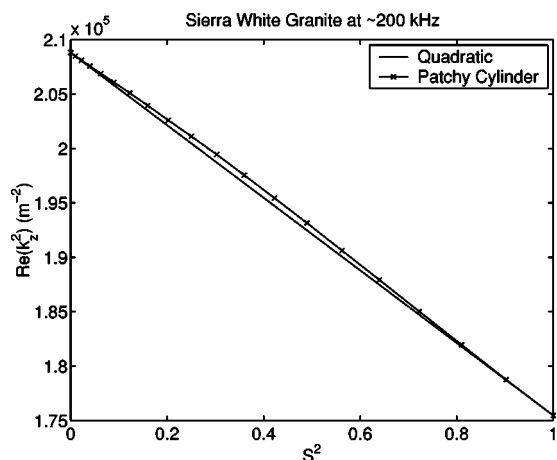


FIG. 10. Comparison between the points that solve the dispersion relation in the complex k_z^2 plane and a simple quadratic fit in water saturation S , for the Sierra White granite measured by Murphy.^{44,45} The fit is also good for this case, but not as good as it was for Massillon sandstone.

- 1 J. F. Claerbout, *Imaging the Earth's Interior* (Blackwell Scientific, Oxford, UK, 1985).
- 2 L. Pochhammer, "Über Fortpflanzungsgeschwindigkeiten kleiner Schwingungen in einem unbegrenzten isotropen Kreiscylinder" (English translation: "On the propagation velocities of small vibrations in an infinite isotropic cylinder"), *J. Reine Angew. Math.* **81**, 324–336 (1876).
- 3 C. Chree, "Longitudinal vibrations of a circular bar," *Q. J. Pure Appl. Math.* **21**, 287–298 (1886).
- 4 M. A. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low-frequency range," *J. Acoust. Soc. Am.* **28**, 168–178 (1956).
- 5 M. A. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range," *J. Acoust. Soc. Am.* **28**, 179–191 (1956).
- 6 G. H. F. Gardner, "Extensional waves in fluid saturated porous cylinders," *J. Acoust. Soc. Am.* **34**, 36–40 (1962).
- 7 D. L. Johnson and S. Kostek, "A limitation of the Biot–Gardner theory of extensional waves in fluid-saturated porous cylinders," *J. Acoust. Soc. Am.* **97**, 741–744 (1995).
- 8 J. G. Berryman, "Dispersion of extensional waves in fluid-saturated porous cylinders at ultrasonic frequencies," *J. Acoust. Soc. Am.* **74**, 1805–1812 (1983).
- 9 J. E. White, "Computed seismic speeds and attenuation in rocks with partial gas saturation," *Geophysics* **40**, 224–232 (1975).
- 10 A. R. Gregory, "Fluid saturation effects in dynamic elastic properties of sedimentary rocks," *Geophysics* **41**, 895–921 (1976).
- 11 W. F. Murphy III, "Effects of partial water saturation on attenuation in Massillon sandstone and Vycor porous-glass," *J. Acoust. Soc. Am.* **71**, 1458–1468 (1982).
- 12 J. G. Berryman, L. Thigpen, and R. C. Y. Chin, "Bulk wave propagation for partially saturated porous solids," *J. Acoust. Soc. Am.* **84**, 360–373 (1988).
- 13 R. Knight and R. Nolen-Hoeksema, "A laboratory study of the dependence of elastic wave velocities on pore scale fluid distribution," *Geophys. Res. Lett.* **17**, 1529–1532 (1990).
- 14 A. N. Norris, "Low-frequency dispersion and attenuation in partially saturated rocks," *J. Acoust. Soc. Am.* **94**, 359–370 (1993).
- 15 R. Knight, J. Dvorkin, and A. Nur, "Acoustic signatures of partial saturation," *Geophysics* **63**, 132–138 (1998).
- 16 D. L. Johnson, "Theory of frequency dependent acoustics in patchy-saturated porous media," *J. Acoust. Soc. Am.* **110**, 682–694 (2001).
- 17 G. Mavko and D. Jizba, "Estimating grain-scale fluid effects on velocity in anisotropic rocks," *Geophysics* **56**, 1940–1949 (1991).
- 18 J. G. Berryman, S. R. Pride, and H. F. Wang, "A differential scheme for elastic properties of rocks with dry or saturated cracks," *Geophys. J. Int.* **151**, 597–611 (2002).
- 19 J. G. Berryman, "Poroelastic shear modulus dependence on pore-fluid properties arising in a model of thin isotropic layers," *Geophys. J. Int.* **157**, 415–425 (2004).
- 20 M. A. Biot, "Mechanics of deformation and acoustic propagation in porous media," *J. Appl. Phys.* **33**, 1482–1498 (1962).
- 21 D. L. Johnson, T. J. Plona, C. Scala, F. Pasierb, and H. Kojima, "Tortuosity and acoustic slow waves," *Phys. Rev. Lett.* **49**, 1840–1844 (1982).
- 22 J. G. Berryman, "Confirmation of Biot's theory," *Appl. Phys. Lett.* **37**, 382–384 (1980).
- 23 F. Gassmann, "Über die Elastizität poröser Medien" (English translation: On the elasticity of porous media), *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich* **96**, 1–23 (1951).
- 24 J. Geerstma, "The effect of fluid pressure decline on volumetric changes of porous rocks," *Trans. AIME* **210**, 331–340 (1957).
- 25 M. A. Biot and D. G. Willis, "The elastic coefficients of the theory of consolidation," *J. Appl. Mech.* **24**, 594–601 (1957).
- 26 J. Geerstma and D. C. Smit, "Some aspects of elastic wave propagation in fluid-saturated porous solids," *Geophysics* **26**, 169–181 (1961).
- 27 R. D. Stoll, "Acoustic waves in saturated sediments," in *Physics of Sound in Marine Sediments*, edited by L. Hampton (Plenum, New York, 1974), pp. 19–39.
- 28 C. Chree, "The equations of an isotropic elastic solid in polar and cylindrical coordinates: Their solution and application," *Trans. Cambridge Philos. Soc.* **14**, 250–369 (1889).
- 29 A. E. H. Love, *A Treatise on the Mathematical Theory of Elasticity* (Dover, New York, 1944), Sec. 201.
- 30 D. Bancroft, "The velocity of longitudinal waves in cylindrical bars," *Phys. Rev.* **59**, 588–593 (1941).

- ³¹H. Deresiewicz and R. Skalak, "On uniqueness in dynamic poroelasticity," *Bull. Seismol. Soc. Am.* **53**, 783–788 (1963).
- ³²J. G. Berryman and L. Thigpen, "Linear dynamic poroelasticity with microstructure for partially saturated porous solids," *J. Appl. Mech.* **52**, 345–350 (1985).
- ³³S. R. Pride and M. W. Haarsen, "Electroseismic wave properties," *J. Acoust. Soc. Am.* **100**, 1301–1315 (1996).
- ³⁴S. R. Pride and E. G. Flekkoy, "Two-phase flow through porous media in the fixed-contact-line regime," *Phys. Rev. E* **60**, 4285–4299 (1999).
- ³⁵P. B. Nagy and G. Blaho, "Experimental measurements of surface stiffness on water-saturated porous solids," *J. Acoust. Soc. Am.* **95**, 828–835 (1994).
- ³⁶P. B. Nagy and A. H. Nayfeh, "Generalized formula for the surface stiffness of fluid-saturated porous media containing parallel pore channels," *Appl. Phys. Lett.* **67**, 1827–1829 (1995).
- ³⁷S. R. Pride, J. G. Berryman, and J. M. Harris, "Seismic attenuation due to wave-induced flow," *J. Geophys. Res.* **109**, B01201 (2004).
- ³⁸J. G. Berryman, "Origin of Gassmann's equations," *Geophysics* **64**, 1627–1629 (1999).
- ³⁹J. G. Berryman, "Modeling high-frequency acoustic velocities in patchy and partially saturated porous rock using differential effective medium theory," *Int. J. Multiscale Computational Engineering* **2**, 115–131 (2004).
- ⁴⁰Y. L. Luke, "Bessel functions of integer order," in *Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun, (Dover, New York, 1965), Chap. 9.
- ⁴¹J. G. Berryman, P. A. Berge, and B. P. Bonner, "Transformation of seismic velocity data to extract porosity and saturation values for rocks," *J. Acoust. Soc. Am.* **107**, 3018–3027 (2000).
- ⁴²J. G. Berryman and H. F. Wang, "Dispersion in poroelastic systems," *Phys. Rev. E* **64**, 011303 (2001).
- ⁴³C. J. Wisse, D. M. J. Smeulders, M. E. H. van Dongen, and G. Chao, "Guided wave modes in porous cylinders: Experimental results," *J. Acoust. Soc. Am.* **112**, 890–895 (2002).
- ⁴⁴W. F. Murphy III, *Effects of microstructure and pore fluids on the acoustic properties of granular sedimentary materials*, Ph.D. thesis, Stanford University, Stanford, CA, 1982.
- ⁴⁵W. F. Murphy III, "Acoustic measures of partial gas saturation in tight sandstones," *J. Geophys. Res.* **89**, 11549–11559 (1984).
- ⁴⁶J. G. Berryman, "Fluid effects on shear waves in finely layered porous media," *Geophysics* (in press).
- ⁴⁷S. R. Pride, E. Tromeur, and J. G. Berryman, "Biot slow-wave effects in stratified rock," *Geophysics* **67**, 271–281 (2002).

A first-order statistical smoothing approximation for the coherent wave field in random porous media

Tobias M. Müller^{a)} and Boris Gurevich^{b)}

Department of Exploration Geophysics, Curtin University^{c)} GPO Box U1987, Perth,
Western Australia 6845, Australia

(Received 6 October 2004; revised 24 November 2004; accepted 26 January 2005)

An important dissipation mechanism for waves in randomly inhomogeneous poroelastic media is the effect of wave-induced fluid flow. In the framework of Biot's theory of poroelasticity, this mechanism can be understood as scattering from fast into slow compressional waves. To describe this conversion scattering effect in poroelastic random media, the dynamic characteristics of the coherent wavefield using the theory of statistical wave propagation are analyzed. In particular, the method of statistical smoothing is applied to Biot's equations of poroelasticity. Within the accuracy of the first-order statistical smoothing an effective wave number of the coherent field, which accounts for the effect of wave-induced flow, is derived. This wave number is complex and involves an integral over the correlation function of the medium's fluctuations. It is shown that the known one-dimensional (1-D) result can be obtained as a special case of the present 3-D theory. The expression for the effective wave number allows to derive a model for elastic attenuation and dispersion due to wave-induced fluid flow. These wavefield attributes are analyzed in a companion paper. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1871754]

PACS numbers: 43.20.Jr, 43.20.Gp, 43.20.-f, 43.20.Wd [JJM]

Pages: 1796–1805

I. INTRODUCTION

Understanding elastic wave attenuation and dispersion in porous fluid-saturated materials is important in such diverse applications as architectural acoustics, soil and rock mechanics, and exploration seismology. One major cause of elastic wave attenuation is viscous dissipation due to the flow of the pore fluid induced by the passing wave. Wave-induced fluid flow occurs as a passing wave creates local pressure gradients within the fluid phase and the resulting fluid flow is accompanied with internal friction until the pore pressure is equilibrated. The fluid flow can take place on various length scales: for example, from compliant fractures into the equant pores (so-called squirt flow^{1–3}), or between mesoscopic heterogeneities like fluid patches in partially saturated rocks.^{4–7}

Theoretical studies of the elastic wave attenuation due to wave-induced flow go back to the 1970s. In such studies wave propagation in an inhomogeneous porous medium is usually analyzed using Biot's equations of poroelasticity with spatially varying coefficients.⁸ The first models of attenuation due to wave-induced flow considered flow caused by a regular assemblage of inhomogeneities of ideal shape such as two concentric spheres or flat slabs.^{9–11} A general theory of wave propagation in heterogeneous porous media using the double-porosity approach was recently developed by Pride and Berryman¹² and Pride *et al.*¹³ In the case of patchy saturation, the results of this theory coincide with those obtained earlier by Johnson¹⁴ using a slightly different method. Although mathematical formulations in the theories Pride and Berryman¹² and Johnson¹⁴ allow for arbitrary geometry of the inhomogeneities, closed-form expressions for

attenuation and phase velocity can only be obtained for a regular system of simple geometrical shapes such as concentric spheres.

In real rocks heterogeneities are more likely to be spatially distributed in a random fashion. Therefore, it is desirable, as suggested by Lopatnikov and Gurevich,¹⁵ to model wave-induced flow using the theory of waves in random media.¹⁶ For one-dimensional (1-D) porous media, Gurevich and Lopatnikov¹⁷ showed that elastic wave attenuation in a randomly layered porous medium differs significantly from attenuation in periodically layered porous media.^{10,18} This suggests that the effects of three-dimensional inhomogeneities on elastic wave wave attenuation and dispersion in porous media may also be different for random and periodic spatial configurations. One approach to this problem is to consider a homogeneous porous medium with randomly distributed discrete, regularly shaped 3-D inclusions. However, so far only the attenuation due to conversion scattering at spherical inclusions from fast P waves into Biot's slow P wave is known.¹⁹

Our purpose in this paper is to analyze the effect of wave-induced fluid flow on the dynamic characteristics of coherent elastic waves propagating in a porous medium whose properties are continuous random functions of position. Our approach is based on the statistical wave theory, as applied to Biot's equations of poroelasticity.⁸ We restrict our analysis to the case of mesoscopic inhomogeneities, that is, inhomogeneities whose characteristic size a is much larger than the typical size of pores or grains a_{pore} but, at the same time, much smaller than the wavelength of the propagating elastic wave, λ :

$$a_{\text{pore}} \ll a \ll \lambda. \quad (1)$$

In other words, we ignore pore-scale heterogeneities, which

^{a)}Electronic mail: tobias.muller@geophy.curtin.edu.au

^{b)}Electronic mail: boris.gurevich@geophy.curtin.edu.au

^{c)}URL: <http://www.geophysics.curtin.edu.au>

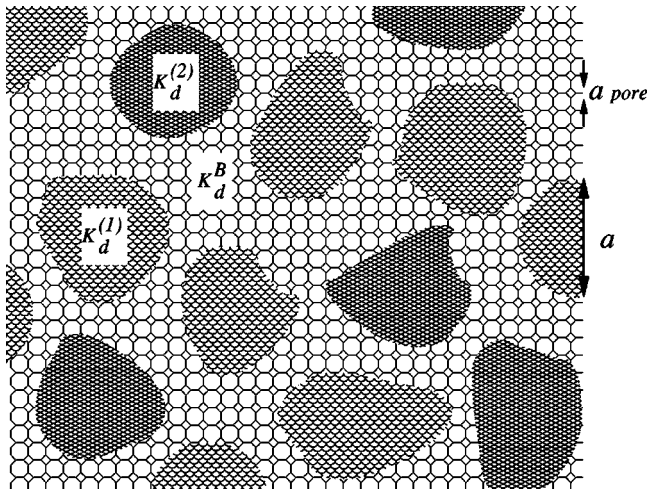


FIG. 1. Sketch of a heterogeneous porous medium where only the bulk modulus of the drained-framework-of-grains fluctuates. The homogeneous background medium is characterized by the bulk modulus K_d^B . There are inhomogeneities with higher bulk moduli, that is $K_d^{(1)} > K_d^B$, and other inhomogeneities with lower bulk modulus, $K_d^{(2)} < K_d^B$. The inhomogeneities have a length scale of a that is much larger than the typical pore scale a_{pore} . During the compression cycle of a wave there will be fluid flow from inhomogeneities with $K_d^{(2)}$ into the background and flow from the background into inhomogeneities with $K_d^{(1)}$. During the extension cycle of the wave the fluid flow becomes reversed.

allows the application of Biot's equations for a poroelastic continuum.⁸ We model wave propagation in the inhomogeneous porous medium using Biot's equations of poroelasticity with randomly varying coefficients (see Fig. 1). By using a Green's function approach, these partial differential equations transform into a system of integral equations. This latter system is solved by means of the method of statistical smoothing which is widely used in problems of electromagnetic, acoustic, and elastic wave propagation.^{20–22} More precisely, we employ a first-order statistical smoothing approximation that sometimes is referred to as a Bourret approximation.^{16,23} For poroelastic media, this method was earlier employed to compute P -wave attenuation and dispersion due to conversion scattering from fast P waves into Biot's slow waves in a randomly layered porous medium.¹⁷ We follow this strategy and analyze the conversion scattering from fast P waves into Biot's slow wave in 3-D randomly inhomogeneous porous media. Biot's slow wave is a highly dissipative wave mode. Therefore, the use of the first-order statistical smoothing approximation to the conversion scattering problem in Biot's equations of poroelasticity quantifies the dissipation of wave field energy due to energy transfer from the coherent component of the fast P wave into the dissipative slow P -wave mode. This is different from the usual application of the method of smoothing to energy conserving wave systems, where an apparent dissipation (so-called scattering attenuation) results from the energy transfer from the coherent component of the wave field into the incoherent component.

In light of previous results for randomly layered porous media, it is useful to introduce some simplifications from the outset. First of all, we restrict our analysis to low frequencies. Specifically, we assume that frequency ω is much

smaller than the critical Biot frequency $\omega_B = \phi \eta / \kappa_0 \rho_f$:

$$\omega \ll \omega_B, \quad (2)$$

where ϕ and κ_0 denote the porosity and permeability of the composite material, while ρ_f and η are the density and viscosity of the pore fluid. Condition (2) implies that the standard Biot's visco-inertial attenuation and dispersion (the so-called Biot's global flow mechanism) is neglected and the P wave number k_p is real. Furthermore, at low frequencies the slow wave is much slower than the fast P wave and therefore the ratio of k_p to slow P wave number k_{ps} is a small parameter:

$$\frac{|k_p|}{|k_{ps}|} \ll 1. \quad (3)$$

Wherever applicable we make use of relation (3), being aware of the underlying low-frequency assumption.

The strategy of this paper is as follows. First, we apply the method of Green's functions to Biot's equations of poroelasticity in order to represent the wavefield due to point source excitation (Sec. II). Next, we derive an integral wavefield representation for the case when the coefficients in Biot's equations exhibit a randomly fluctuating component. This wavefield representation (also called the scattering equation) is then converted into an integral equation for Green's function for the inhomogeneous medium. In Sec. III we apply a perturbation method to compute an approximation for the mean of the Green's function. From this Green's function we derive and analyze an explicit expression for the effective, complex P wave number that accounts for the conversion scattering from P into slow P waves. The range of applicability of our theory is analyzed in Sec. IV. Finally, in Sec. V we show how the results of the aforementioned 1-D theory can be recovered as a limiting case of the more general 3-D theory. A detailed analysis of attenuation and dispersion as functions of frequency and the potential applicability of these wavefield signatures to interpret the effect of wave-induced flow in real rocks is presented in a companion paper.²⁴

II. FORMULATION OF THE POROELASTIC SCATTERING PROBLEM

A. Green's function approach for Biot's equations of poroelasticity

In order to study dynamic effects of elastic wave propagation in porous media, we base our analysis on Biot's equation of poroelasticity.⁸ Using index notation—summation over repeated indices is assumed and partial derivatives are denoted as $_{,i}$ or ∂_i —we can write the equations of motion in the frequency domain (the time-harmonic dependency $\exp(-i\omega t)$ is omitted),

$$\rho \omega^2 u_i + \rho_f \omega^2 w_i + \tau_{ij,j} = 0, \quad (4)$$

$$\rho_f \omega^2 u_i + q \omega^2 w_i - p_{,i} = 0, \quad (5)$$

where τ_{ij} is the total stress tensor, p the fluid pressure, while u_i and w_i are the components of the solid and relative fluid displacement vectors, respectively. The relative fluid displacement is defined as $w_i = \phi(U_i - u_i)$, where is U_i the

fluid displacement. The densities of the solid and fluid phase are denoted by ρ_g and ρ_f so that the bulk density is given by $\rho = \phi\rho_f + (1 - \phi)\rho_g$, where ϕ is porosity. The parameter q is defined as $q = i\eta/(\omega\kappa_0)$, where η is viscosity and κ_0 permeability. We note that this definition of q is a consequence of the low-frequency assumption (2).

In order to obtain a closed system of wave equations in the displacements u_i and w_i , we complement the equations of motion with the stress-strain relations for an isotropic poroelastic medium,⁸

$$\tau_{ij} = G[u_{i,j} + u_{j,i} - 2\delta_{ij}u_{j,j}] + \delta_{ij}[Hu_{j,j} + Cw_{j,j}], \quad (6)$$

$$p = -Cu_{j,j} - Mw_{j,j}. \quad (7)$$

Here G is the porous-material shear modulus, and H is the undrained, low-frequency P -wave modulus given by Gassmann's equation:

$$H = P_d + \alpha^2 M, \quad (8)$$

where

$$M = [(\alpha - \phi)/K_g + \phi/K_f]^{-1}. \quad (9)$$

In Eqs. (8)–(9), $P_d = K_d + 4/3G$ is the P -wave modulus of the drained frame, $\alpha = 1 - K_d/K_g$ is the Biot–Willis coefficient, $C = \alpha M$, and K_g , K_d , and K_f denote the bulk moduli of the solid phase, the drained frame, and the fluid phase, respectively. Symbol δ_{ij} is Kronecker's delta (the identity tensor).

It is expedient to write the above system of coupled wave equations in matrix form:

$$\begin{bmatrix} L_{ik}^{(1)} & L_{ik}^{(2)} \\ L_{ik}^{(3)} & L_{ik}^{(4)} \end{bmatrix} \cdot \begin{bmatrix} u_k \\ w_k \end{bmatrix} = \mathbf{0}, \quad (10)$$

where we defined the linear differential operators as follows:

$$L_{ik}^{(1)} = \rho\omega^2\delta_{ik} + \partial_j G[\delta_{jk}\partial_i + \delta_{ik}\partial_j - 2\delta_{ij}\partial_k] + \partial_i H\partial_k, \quad (11)$$

$$L_{ik}^{(2)} = \rho_f\omega^2\delta_{ik} + \partial_i C\partial_k, \quad (12)$$

$$L_{ik}^{(3)} = L_{ik}^{(2)}, \quad (13)$$

$$L_{ik}^{(4)} = q\omega^2\delta_{ik} + \partial_i M\partial_k. \quad (14)$$

Note that in inhomogeneous media all the poroelastic parameters are functions of position. In the presence of point sources, the right-hand side of Eq. (10) can be written as

$$-\begin{bmatrix} F_i^0 \delta(r_i - r'_i) \\ f_i^0 \delta(r_i - r'_i) \end{bmatrix}, \quad (15)$$

where F_k^0 and f_k^0 represent constant forces applied to the bulk and fluid phase, respectively, and $\delta(r_i - r'_i)$ denotes the Dirac delta function. The response of system (10) to point sources of the form (15) can be formulated as²⁵

$$\begin{bmatrix} u_i \\ w_i \end{bmatrix} = \begin{bmatrix} G_{ik}^F & G_{ik}^f \\ G_{ik}^f & G_{ik}^w \end{bmatrix} \cdot \begin{bmatrix} F_k^0 \\ f_k^0 \end{bmatrix}, \quad (16)$$

where G_{ik}^F , G_{ik}^f , and G_{ik}^w denote the Green's tensors. Thus, the point source response of system (10) in an isotropic unbounded medium is described by three Green's tensors. Explicit expressions for the G_{ik} 's for a homogeneous medium

are given by Pride and Haartsen²⁵ and reproduced in Appendix A. The wavefields observed at position \mathbf{r} due to arbitrary point sources F_i and f_i , applied at position \mathbf{r}' , can be expressed by a convolution equation of the form

$$\begin{bmatrix} u_i(\mathbf{r}) \\ w_i(\mathbf{r}) \end{bmatrix} = \int_{V'} d^3\mathbf{r}' \begin{bmatrix} G_{ik}^F(\mathbf{r} - \mathbf{r}') & G_{ik}^f(\mathbf{r} - \mathbf{r}') \\ G_{ik}^f(\mathbf{r} - \mathbf{r}') & G_{ik}^w(\mathbf{r} - \mathbf{r}') \end{bmatrix} \cdot \begin{bmatrix} F_k(\mathbf{r}') \\ f_k(\mathbf{r}') \end{bmatrix}. \quad (17)$$

On the basis of Eq. (17), we now derive a wavefield representation in a randomly inhomogeneous medium.

B. The basic scattering equation

In randomly inhomogeneous porous media, all poroelastic parameters can be presented as random fields $X(\mathbf{r})$. To be more specific, we assume that each of these poroelastic parameters is the sum of a constant background value, \bar{X} , and a fluctuating part, $\tilde{X}(\mathbf{r})$, so that

$$X = \bar{X} + \tilde{X} = \bar{X}(1 + \varepsilon_X), \quad (18)$$

where $\varepsilon_X = \tilde{X}/\bar{X}$ denotes the relative fluctuations. The average over the ensemble of the realizations (denoted by $\langle \cdot \rangle$) of ε_X is assumed to be zero: $\langle \varepsilon_X \rangle = 0$. The spatial correlation function of two random fields is defined as

$$B_{XX}(\delta\mathbf{r}) = \langle \varepsilon_X(\mathbf{r} + \delta\mathbf{r})\varepsilon_X(\mathbf{r}) \rangle, \quad (19)$$

where the dependence of B on the difference vector $\delta\mathbf{r}$ only is a consequence of the assumption of statistically homogeneous random fields,¹⁶ which we use throughout this paper. The variance of the random process ε_X will be denoted as $B_{XX}(0) = \langle \varepsilon_X^2 \rangle = \sigma_{XX}^2$. Using (18), the differential operators L_{ik} can be also decomposed as²¹

$$L_{ik} = \bar{L}_{ik} + \tilde{L}_{ik}, \quad (20)$$

where the perturbing operator \tilde{L}_{ik} satisfies $\langle \tilde{L}_{ik} \rangle = 0$. As is typical for statistical wave problems, in the following we assume that the constant part \bar{X} and statistical properties of the fluctuations ε_X are known. Therefore, rather than seeking the solution of Eq. (17) for a given realization ε_X , we seek statistical moments of the solution for given statistical properties of fluctuations.

The substitution of (20) into matrix equation (10) yields

$$\begin{bmatrix} \bar{L}_{ik}^{(1)} & \bar{L}_{ik}^{(2)} \\ \bar{L}_{ik}^{(2)} & \bar{L}_{ik}^{(4)} \end{bmatrix} \cdot \begin{bmatrix} u_k \\ w_k \end{bmatrix} = - \begin{bmatrix} \tilde{L}_{ik}^{(1)} & \tilde{L}_{ik}^{(2)} \\ \tilde{L}_{ik}^{(2)} & \tilde{L}_{ik}^{(4)} \end{bmatrix} \cdot \begin{bmatrix} u_k \\ w_k \end{bmatrix}, \quad (21)$$

In the most general case, the perturbing operators \tilde{L}_{ik} contain fluctuations of all poroelastic moduli and densities. The right-hand side of (21) can be thought of as a source term in the homogeneous system (10) due to the presence of inhomogeneities (so-called secondary sources). Thus, Eq. (21) can be understood as an inhomogeneous equation with constant coefficients, whose formal solution can be written by substituting the source term into Eq. (17):

$$\begin{bmatrix} u_i \\ w_i \end{bmatrix} = \begin{bmatrix} u_i^0 \\ w_i^0 \end{bmatrix} + \int_V dV \begin{bmatrix} G_{ij}^F & G_{ij}^f \\ G_{ij}^f & G_{ij}^w \end{bmatrix} \cdot \begin{bmatrix} \tilde{L}_{jk}^{(1)} & \tilde{L}_{jk}^{(2)} \\ \tilde{L}_{jk}^{(2)} & \tilde{L}_{jk}^{(4)} \end{bmatrix} \cdot \begin{bmatrix} u_k \\ w_k \end{bmatrix}. \quad (22)$$

Equation (22) is the basic poroelastic scattering equation. The total wavefields u_i and w_i are composed of wavefields propagating in the homogeneous background medium, u_i^0 and w_i^0 , and scattered wavefields (the second term). By definition, u_i^0 and w_i^0 satisfy the homogeneous equation (10). The scattered wavefields are represented by volume integrals whose kernels involve the Green's tensors and the secondary sources. The scattered wavefields vanish if there are no fluctuations in the medium parameters. The integration volume encompasses the inhomogeneous part of the medium, which in our case is the whole 3-D space. We note that Eq. (22) is closely related to the representation integral of the scattered field in an unbounded medium (for the poroelastic formulation of this theorem we refer to Norris²⁶ and Pride and Haartsen²⁵). According to Eq. (17), the wavefields can be represented as a convolution of Green's tensors with the source function. Let us denote Green's tensors for the homogeneous background medium by ${}^0G_{ik}^{F,f,w}$ and for the inhomogeneous medium by $G_{ik}^{F,f,w}$. Substituting these wavefield representations into Eq. (22), we obtain an equation for the Green's tensors of the inhomogeneous medium,

$$\begin{bmatrix} G_{im}^F & G_{im}^f \\ G_{im}^f & G_{im}^w \end{bmatrix} = \begin{bmatrix} {}^0G_{im}^F & {}^0G_{im}^f \\ {}^0G_{im}^f & {}^0G_{im}^w \end{bmatrix} + \int_V dV \begin{bmatrix} {}^0G_{ij}^F & {}^0G_{ij}^f \\ {}^0G_{ij}^f & {}^0G_{ij}^w \end{bmatrix} \cdot \begin{bmatrix} \tilde{L}_{jk}^{(1)} & \tilde{L}_{jk}^{(2)} \\ \tilde{L}_{jk}^{(2)} & \tilde{L}_{jk}^{(4)} \end{bmatrix} \cdot \begin{bmatrix} G_{km}^F & G_{km}^f \\ G_{km}^f & G_{km}^w \end{bmatrix}. \quad (23)$$

In order to simplify the equations that follow, we introduce a shorthand notation. The latter equation can be symbolically rewritten as

$$\mathbf{G} = \mathbf{G}^0 + \int \mathbf{G}^0 \tilde{\mathbf{L}} \mathbf{G}, \quad (24)$$

where \mathbf{G} , \mathbf{G}^0 , and $\tilde{\mathbf{L}}$ represent matrices, whose elements are tensors of rank two, and matrix multiplication rules apply. In (24) we also omitted the integration volume for brevity.

III. FIRST-ORDER STATISTICAL SMOOTHING OF BIOT'S EQUATIONS OF POROELASTICITY

A. Mean Green's tensor

We will now analyze Eq. (24) using a statistical approach. Since the matrix of perturbing operators $\tilde{\mathbf{L}}$ in Eq. (24) contains fluctuating medium parameters, the resulting matrix of Green's tensors also contains randomly fluctuating elements. Because individual realizations of the random wavefields are never known, it is natural to analyze the statistical moments of \mathbf{G} . Solving Eq. (24) by iteration we obtain the scattering series

$$\mathbf{G} = \mathbf{G}^0 + \int \mathbf{G}^0 \tilde{\mathbf{L}} \mathbf{G}^0 + \int \int \mathbf{G}^0 \tilde{\mathbf{L}} \mathbf{G}^0 \tilde{\mathbf{L}} \mathbf{G}^0 + \int \int \int \dots \quad (25)$$

Averaging this equation by the ensemble of realizations and regrouping the scattering terms yields

$$\bar{\mathbf{G}} = \mathbf{G}^0 + \int \int \mathbf{G}^0 \mathbf{Q} \bar{\mathbf{G}}, \quad (26)$$

where $\bar{\mathbf{G}} = \langle \mathbf{G} \rangle$ is the matrix of mean Green's tensors, and \mathbf{Q} is the matrix operator defined as

$$\mathbf{Q} = \begin{bmatrix} Q_{ik}^{(1)} & Q_{ik}^{(2)} \\ Q_{ik}^{(3)} & Q_{ik}^{(4)} \end{bmatrix} = \left\langle \tilde{\mathbf{L}} \mathbf{G}^0 \tilde{\mathbf{L}} + \int \tilde{\mathbf{L}} \mathbf{G}^0 \tilde{\mathbf{L}} \mathbf{G}^0 \tilde{\mathbf{L}} + \int \dots \right\rangle. \quad (27)$$

Operator \mathbf{Q} given by Eq. (27) corresponds to the kernel-of-mass operator in the acoustic formulation.¹⁶ The linear integral equation in $\bar{\mathbf{G}}$ [Eq. (26)] is the poroelastic analog of the Dyson equation (see also Gurevich and Lopatnikov,²⁷ where an analogous equation for the mean field is derived). It is not possible to obtain an exact solution of Eq. (26). A first-order statistical smoothing consists in the first-order truncation of the infinite series expression for the operator \mathbf{Q} . Then, we obtain the following approximation for the mean Green's tensor:

$$\bar{\mathbf{G}} = \mathbf{G}^0 + \int \int \mathbf{G}^0 \langle \tilde{\mathbf{L}} \mathbf{G}^0 \tilde{\mathbf{L}} \rangle \bar{\mathbf{G}} \quad (28)$$

$$= \mathbf{G}^0 + \int \int \mathbf{G}^0 \mathbf{Q}^B \bar{\mathbf{G}}. \quad (29)$$

The truncation of the series (27) implies that the first-order statistical smoothing is valid when $|\varepsilon_X| \ll 1$, i.e., when the absolute value of the relative fluctuations of X is a small parameter. Note also that the elements of matrix operator \mathbf{Q}^B only contain terms involving the second statistical moment of the fluctuating parts of the \tilde{L}_{ik} 's, that is, they are of the order $O(\varepsilon^2)$. Higher-order correlations are neglected within the accuracy of the first-order statistical smoothing approximation.

Since Eq. (29) contains a double volume convolution, it is expedient to work with its spatial Fourier transform:

$$\bar{\mathbf{g}} = \mathbf{g}^0 + (8\pi^3)^2 \mathbf{g}^0 \mathbf{q} \bar{\mathbf{g}}, \quad (30)$$

where $\bar{\mathbf{g}}$, \mathbf{g}^0 , and \mathbf{q} denote the spatial Fourier transforms of $\bar{\mathbf{G}}$, \mathbf{G}^0 , and \mathbf{Q}^B , respectively (see Appendix A for a definition of the Fourier transform pair). Equation (30) is of the same form as the equation for the mean Green's function in acoustic random media. However, in contrast to the acoustic case, Eq. (30) is not a simple algebraic equation for the mean Green's tensors, but a system of four tensorial equations for the three unknown mean Green's tensors. Formally, we can express the solution of (30) in the form $\bar{\mathbf{g}} = \mathbf{W}^{-1} \mathbf{g}^0$, where \mathbf{W}^{-1} is the inverse of the matrix $\mathbf{W} = \mathbf{I} - (8\pi^3)^2 \mathbf{g}^0 \mathbf{q}$ with the identity tensor \mathbf{I} . However, the computation of \mathbf{W}^{-1} is cumbersome. Instead, we are looking for a more feasible way to compute some elements of $\bar{\mathbf{g}}$. Carrying out the necessary matrix multiplications in (30), we find that this system splits up into two pairs of coupled equations. Since we are only interested in the characteristics of the fast P wave, which are exclusively contained in the Green's tensor \bar{g}^F [see also Eqs. (16) and (A1)], we analyze only those two equations that involve \bar{g}_{ik}^F . We obtain

$$\begin{aligned} \bar{g}^F = g^F + (8\pi^3)^2 [g^F q^{(1)} \bar{g}^F + g^F q^{(2)} \bar{g}^f + g^f q^{(3)} \bar{g}^F \\ + g^f q^{(4)} \bar{g}^f], \end{aligned} \quad (31)$$

$$\begin{aligned} \bar{g}^f = g^f + (8\pi^3)^2 [g^f q^{(1)} \bar{g}^F + g^f q^{(2)} \bar{g}^f + g^w q^{(3)} \bar{g}^F \\ + g^w q^{(4)} \bar{g}^f], \end{aligned} \quad (32)$$

where we omitted subscripts for brevity. The quantities g without an upper bar denote the background space Green's tensors. Since all quantities $q^{(i)}$ ($i=1, \dots, 4$) are of the order $O(\varepsilon^2)$, \bar{g}^f is also of the order $O(\varepsilon^2)$. Inserting the expression for \bar{g}^f [Eq. (32)] into Eq. (31) and neglecting terms of higher order than $O(\varepsilon^2)$, we obtain

$$\begin{aligned} \bar{g}^F = g^F + (8\pi^3)^2 [q^F q^{(1)} \bar{g}^F + g^F q^{(2)} g^f + g^f q^{(3)} \bar{g}^F \\ + g^f q^{(4)} g^f]. \end{aligned} \quad (33)$$

Equation (33) is an implicit equation for the mean Green's tensor \bar{g}^F . Because of its tensorial character, an explicit solution for \bar{g}^F is still difficult to construct. Note, however, that we are not interested in the mean Green's tensor itself but only in a mean P wave number contained in \bar{g}^F .

B. Effective wave number of the fast P -wave

In order to extract an effective wave number from Eq. (33) we have to introduce further simplifications. Because of the assumption of small fluctuations in the medium parameters ($\varepsilon \ll 1$), we can expect that the fluctuations of the wavefield are also small if the wavelengths are much larger than the size of the inhomogeneities. Then, we can assume that mean Green's tensor $\bar{g}_{ik}^F(\mathbf{K})$ is of the same functional form as a background Green's tensor $g_{ik}^F(\mathbf{K})$ given by Eq. (A8), however, involving some effective P -wave number (and also effective bulk density).

Let us construct a simple case, where most of the Green's tensor components vanish. This can be achieved using the following procedure: We consider an incoming, plane P wave propagating in the x_3 direction (i.e., only the displacement component u_3 is nonzero). The resulting coherent P wave in the inhomogeneous medium will also propagate in the x_3 direction. Therefore, only the tensor components $i=j=3$ of g_{ij}^F need to be analyzed. Noting that in this case the Green's tensor $g_{ik}^F(\mathbf{K})$ yields the largest contribution for the spatial wave number $K=k_p$, we can approximate the full Green's tensor (A8) by

$$g_{33}^F \approx \frac{-1}{8\pi^3 \rho \omega^2} \left(1 + \frac{K^2}{k_p^2 - K^2} \right). \quad (34)$$

We assume that the mean Green's tensor component is given by

$$\bar{g}_{33}^F \approx \frac{-1}{8\pi^3 \bar{\rho} \omega^2} \left(1 + \frac{K^2}{\bar{k}_p^2 - K^2} \right), \quad (35)$$

where \bar{k}_p is the searched-for effective P wave number. Substituting Eq. (34) for g^F and Eq. (35) for \bar{g}^F into Eq. (33) we obtain, after algebraic manipulations,

$$\bar{k}_p \approx k_p \left(1 + \frac{4\pi^3}{\rho \omega^2} q_{33}^{(1)} \right). \quad (36)$$

Here we neglected terms that contain combinations of the tensor components $q_{33}^{(i)}$. This introduces no additional inac-

curacy because higher-order correlations are neglected within the accuracy of the first-order statistical smoothing $O(\varepsilon^2)$.

The remaining problem is the evaluation of $q_{33}^{(1)}$ in (36), or equivalently, of $Q_{33}^{(1)}$ in space domain. In explicit form, from the first term in the expansion of \mathbf{Q} as given by Eq. (27), we obtain

$$\begin{aligned} Q_{ik}^{(1)}(\mathbf{r}' - \mathbf{r}'') = & \langle \tilde{L}_{ij}^{(1)}(\mathbf{r}') G_{jl}^F(\mathbf{r}' - \mathbf{r}'') \tilde{L}_{lk}^{(1)}(\mathbf{r}'') \\ & + 2\tilde{L}_{ij}^{(1)}(\mathbf{r}') G_{jl}^f(\mathbf{r}' - \mathbf{r}'') \tilde{L}_{lk}^{(2)}(\mathbf{r}'') \\ & + \tilde{L}_{ij}^{(2)}(\mathbf{r}') G_{jl}^w(\mathbf{r}' - \mathbf{r}'') \tilde{L}_{lk}^{(2)}(\mathbf{r}'') \rangle, \end{aligned} \quad (37)$$

where for statistically homogeneous random media both Q_{ik} and G_{ik} depend only on the difference vector $\mathbf{r}' - \mathbf{r}''$. It is interesting to note that in the elastic limit, only the first term of $Q_{ik}^{(1)}$ is nonzero. In the poroelastic case we have to analyze all three terms. Expression (37) involves the perturbing operators $\tilde{L}_{ij}^{(1)}$ and $\tilde{L}_{ij}^{(2)}$ (but not $\tilde{L}_{ij}^{(4)}$). Let us now specify the perturbing operators resulting from (11) and (12),

$$\tilde{L}_{ik}^{(1)} = \tilde{\rho} \omega^2 \delta_{ik} + \partial_j \tilde{G} [\delta_{jk} \partial_i + \delta_{ik} \partial_j - 2\delta_{ij} \partial_k] + \partial_i \tilde{H} \partial_k, \quad (38)$$

$$\tilde{L}_{ik}^{(2)} = \tilde{\rho}_f \omega^2 \delta_{ik} + \partial_i \tilde{C} \partial_k, \quad (39)$$

for a particular situation in which we can find an explicit analytical expression for the right-hand side of Eq. (37). In the following, we neglect fluctuations of the densities ρ and ρ_f . This is possible because of the restriction to low frequencies. It can be shown that incorporation of density fluctuations yields a correction to the background P wave number [the second term in Eq. (36)], which scales with ω^3 , whereas the other fluctuations result in a ω^2 dependence, as shown below. This simplification is also in accordance with the 1-D result,¹⁷ where the density fluctuations do not appear in the final expression for the effective P wave number though these fluctuations were not neglected *a priori*. These simplifications yield

$$\tilde{L}_{ik}^{(1)} = \partial_k \tilde{G} \partial_i + \partial_j \delta_{ik} \tilde{G} \partial_j - 2\partial_i \tilde{G} \partial_k + \partial_i \tilde{H} \partial_k, \quad (40)$$

$$\tilde{L}_{ik}^{(2)} = \partial_i \tilde{C} \partial_k. \quad (41)$$

A detailed computation of the three $Q_{ik}^{(1)}$ terms in Eq. (37) using the perturbing operators (40) and (41) is provided in Appendix B. The result in the wave number domain can be represented as

$$q_{33}^{(1)} = q^{HH} + q^{HG} + q^{HC} + q^{GG} + q^{GC} + q^{CC}, \quad (42)$$

where

$$\begin{aligned} q^{HH} = & \frac{1}{8\pi^3} k_p^2 \left(\frac{H^2}{P_d} B_{HH}(0) \right. \\ & \left. + \frac{C^2}{N} k_{ps}^2 \int_0^\infty r B_{HH}(r) \exp[ik_{ps} r] dr \right), \end{aligned} \quad (43)$$

$$q^{HG} = -\frac{1}{3\pi^3} k_p^2 \left(\frac{GH}{P_d} B_{HG}(0) + \frac{\alpha^2 MG}{P_d} k_{ps}^2 \int_0^\infty r B_{PM}(r) \exp[ik_{ps}r] dr \right), \quad (44)$$

$$q^{HC} = -\frac{1}{4\pi^3} \frac{C^2}{N} k_p^2 \left(B_{HC}(0) + k_{ps}^2 \int_0^\infty r B_{HC}(r) \exp[ik_{ps}r] dr \right), \quad (45)$$

$$q^{GG} = \frac{1}{15\pi^3} \frac{G}{NH^2} k_p^2 \left([4C^2G + 4NHG + NH^2] B_{GG}(0) + 4C^2G k_{ps}^2 \int_0^\infty r B_{GG}(r) \exp[ik_{ps}r] dr \right) \quad (46)$$

$$q^{GC} = \frac{1}{3\pi^3} \frac{\alpha^2 MG}{P_d} k_p^2 \left(B_{GC}(0) + k_{ps}^2 \int_0^\infty r B_{GC}(r) \exp[ik_{ps}r] dr \right), \quad (47)$$

$$q^{CC} = +\frac{1}{8\pi^3} \frac{C^2}{N} k_p^2 \left(B_{CC}(0) + k_{ps}^2 \int_0^\infty r B_{CC}(r) \exp[ik_{ps}r] dr \right). \quad (48)$$

Here, B_{HH} , B_{HC} , B_{HG} , B_{GG} , B_{GC} , B_{CC} denote the (cross-) correlation functions of the random fields \tilde{H} , \tilde{G} , and \tilde{C} defined by Eq. (19). In the derivation of the $Q_{ik}^{(1)}$ terms we assumed that the random media realizations are statistically isotropic and therefore the correlation functions depend only on r . The upper bar denoting the background properties is omitted. It is important to note that Eqs. (43)–(48) will provide a correction term to the background P wave number [see Eq. (36)], which exclusively accounts for conversion scattering into Biot's slow wave. The separation of terms in the operator $Q_{ik}^{(1)}$ that describe the scattering process from fast into Biot's slow P wave and all other scattering processes is possible because of the low-frequency assumption [Eq. (3)]. In other words, because of the large separation between the two characteristic frequencies for ordinary elastic scattering and conversion scattering into Biot's slow wave, we can distinguish between the different scattering processes, using the frequency dependency of the corresponding terms in the scattering equation (see also Appendix B).

We will now assume that all correlation functions are of the same functional form and only differ by their variances, i.e., $B_{XY} = \sigma_{XY}^2 B(r)$ with $B(0) = 1$ and $B(\infty) = 0$. Substituting then expressions (43)–(48) into Eq. (36) we obtain the final result for the effective P wave number,

$$\bar{k}_p = k_p \left(1 + \Delta_2 + \Delta_1 k_{ps}^2 \int_0^\infty r B(r) \exp[ik_{ps}r] dr \right), \quad (49)$$

with the dimensionless coefficients

$$\Delta_1 = \frac{\alpha^2 M}{2P_d} \left(\sigma_{HH}^2 - 2\sigma_{HC}^2 + \sigma_{CC}^2 + \frac{32}{15} \frac{G^2}{H^2} \sigma_{GG}^2 - \frac{8}{3} \frac{G^2}{H} \sigma_{HG}^2 + \frac{8}{3} \frac{G}{H} \sigma_{GC}^2 \right), \quad (50)$$

$$= \frac{\alpha^2 M}{2P_d} \left(\left(\left(\varepsilon_H - \frac{4}{3} \frac{G}{H} \varepsilon_G - \varepsilon_C \right)^2 \right) + \frac{16}{45} \frac{G^2}{H^2} \sigma_{GG}^2 \right), \quad (51)$$

$$\Delta_2 = \Delta_1 + \frac{1}{2} \sigma_{HH}^2 - \frac{4}{3} \frac{G}{H} \sigma_{HG}^2 + \left(\frac{4G}{H} + 1 \right) \frac{4}{15} \frac{G}{H} \sigma_{GC}^2. \quad (52)$$

The structure of the effective P wave number can be explained as follows: Due to the presence of random inhomogeneities, there are two terms added to the background wave number k_p . The first term, Δ_2 , is frequency independent and consists in a weighted sum of the variances of the random fields \tilde{H} , \tilde{G} , and \tilde{C} . The second term is frequency dependent and contains an integral over the correlation function multiplied by a weighted sum of the variances, Δ_1 . It is important to note that the expression for \bar{k}_p describes only the process of conversion scattering from fast into slow P waves. The contribution of purely elastic scattering is left out. The corresponding result would include additional terms involving the correlation functions B_{HH} , B_{GG} , and B_{HG} that describe the elastic scattering (P to P and S waves) and produces the typical Rayleigh frequency dependence for attenuation.²³ In other words, the second and third terms of the effective P wave number (49) correspond to the mechanism of wave-induced fluid flow only. Therefore, an analysis of the properties of \bar{k}_p gives insight into the relationship between the properties of elastic waves and wave-induced flow. By definition, the real part of \bar{k}_p is related to the phase velocity v through $v(\omega) = \omega / \Re\{\bar{k}_p\}$, whereas the imaginary part yields the attenuation coefficient γ : $\gamma(\omega) = \Im\{\bar{k}_p\}$. From the structure of Eq. (49), it can be seen that the phase velocity of the coherent wave in an equivalent medium is smaller than in the background medium. By inspection we find that $\gamma > 0$, that is, the coherent wave is exponentially damped. A detailed analysis of attenuation and velocity dispersion in an equivalent medium is presented in a companion paper.

IV. RANGE OF APPLICABILITY

The derivation of the dispersion relation (49) and the associated results on attenuation and phase velocity dispersion are based on several assumptions that restrict their range of applicability. The main restriction on our results is due to the use of the first-order statistical smoothing approximation. In the acoustic case the applicability condition of this approximation can be written as¹⁶

$$\sigma_n^2 (ka)^2 \ll 1, \quad (53)$$

where σ_n^2 denotes the variance of the velocity fluctuations, k is the wave number of a wave propagating in the homogeneous background, and a is a characteristic length scale associated with the size of the inhomogeneities. Condition (53) ensures that the correction terms to the background wave

number [similar to those in Eq. (49)] are small. More precisely, this condition was obtained by analyzing the next term in the equation for the kernel-of-mass operator \mathcal{Q} and requiring that the difference between the correction terms from the first-order statistical smoothing and that of the next, higher-order approximation is small. It is natural to assume that a similar condition must hold in the poroelastic case.

Using the first-order statistical smoothing the kernel-of-mass operator \mathbf{Q} can be approximated by $\mathbf{Q}^B \approx \langle \tilde{\mathbf{L}} \mathbf{G}^0 \tilde{\mathbf{L}} \rangle$ [Eq. (29)]. The next term in the infinite series for \mathbf{Q} reads as

$$\mathbf{Q}^{(2)} = \langle \tilde{\mathbf{L}} \mathbf{G} \tilde{\mathbf{L}} \rangle, \quad (54)$$

where $\tilde{\mathbf{G}}$ is the matrix of the mean Green's tensors found from the first-order statistical smoothing. To compute $\mathbf{Q}^{(2)}$, we assume that there is no multiple scatter of slow P waves into slow P waves and, therefore, the mean Green's tensor component \tilde{g}_{33}^F is still of the form (35). Then we can derive from Eq. (33) an equation for the effective P wave number similar to (36), but now involving $q_{33}^{(1)}$ determined from (54). The computation of $q_{33}^{(1)}$ according to (54) is analogous to that shown in the previous section and results in an effective P wave number \bar{k}_{p2} :

$$\bar{k}_{p2} \approx k_p + \Delta_2 k_p + \Delta_1 \bar{k}_p k_{ps}^2 \int_0^\infty r B(r) \exp[ik_{ps} r] dr, \quad (55)$$

with \bar{k}_p given by (49). Let us compare this result with the result for the effective wave number using first-order statistical smoothing [Eq. (49)]. Assuming that \bar{k}_p can be represented as a sum of background wave number k_p and a correction term Δk , the difference between (55) and (49) becomes

$$\bar{k}_{p2} - \bar{k}_p = \Delta k \Delta_1 k_{ps}^2 \int_0^\infty r B(r) \exp[ik_{ps} r] dr. \quad (56)$$

A necessary condition for the significance of the correction term using the first-order statistical smoothing Δk is the smallness of the difference (56) compared to Δk :

$$\left| \frac{\bar{k}_{p2} - \bar{k}_p}{\Delta k} \right| \ll 1. \quad (57)$$

Using for instance the correlation function $B(r) = \exp(-|r|/a)$, we obtain, from (56) and (57),

$$\max\{\Delta_1 (|k_{ps}|a)^2, \Delta_2\} \ll 1. \quad (58)$$

Relation (58) gives an estimate of the applicability of the first-order statistical smoothing approximation. That is, our results can be used in the case of weak-contrast media and weak wavefield fluctuations.

Physically, the existence of slow P waves is associated with the equilibration of pore pressure that at low frequencies is controlled by the diffusion equation with diffusion length¹⁸ $\lambda_D = \sqrt{\kappa_0 N / \omega \eta}$. Therefore, the interplay between λ_D and the correlation length a defines two different regimes. If $\lambda_D > a$ then the wave-induced pressure disturbance is equilibrated. This relation holds for low frequencies and thus defines the low-frequency or relaxed regime. Conversely, if $\lambda_D < a$ then there is not enough time for the pore

pressure to relax. This unrelaxed case occurs at high frequencies. A characteristic frequency that separates these low- and high-frequency regimes can be written as

$$\omega_c = \frac{\kappa_0 N}{a^2 \eta}. \quad (59)$$

Note that this characteristic frequency is identical to the frequency where maximal attenuation occurs.²⁴ The existence of both regimes within our wavefield approximation, which is based on the low-frequency approximation of Biot's equations, is only possible if

$$\omega_c \ll \omega_B, \quad (60)$$

which imposes the additional condition for the average size of the mesoscopic inhomogeneities,

$$a^2 \gg \frac{\kappa_0 N}{\eta \omega_B}. \quad (61)$$

Conditions (58) and (61) define the range of applicability of our results. Whereas condition (60) is a necessary condition for the validity of our results, there exists another condition for the observability of wave-induced flow. Only if ω_c is sufficiently different from the characteristic frequency, where elastic scattering of the fast wave modes dominates (i.e., $P \rightarrow P$ and $P \rightarrow S$), it is possible to distinguish between the two processes. Elastic scattering will dominate at frequency $\omega_S = c/a$, where c is the phase velocity of the fast P -wave mode. Therefore, the observability condition $\omega_S \gg \omega_c$ imposes $a^2 \gg \kappa_0^2 N \rho / (\eta^2 H)$.

V. REDUCTION TO THE 1-D CASE

In order to further substantiate our results, we analyze their connection with the known 1-D result. For a system of randomly layered porous media an effective, complex P wave number was obtained earlier by Gurevich and Lopatnikov;¹⁷ see their equation (56). This 1-D result is also based on the method of (first-order) statistical smoothing. If only the parameter M fluctuates, the 1-D result can be rewritten as

$$\bar{k}_p^{1D} = k_p^* \left[1 + \frac{1}{2} i \frac{\alpha^2 M}{P_d} k_{ps}^* \int_0^\infty dz B_{MM}(z) \exp[ik_{ps}^* z] \right], \quad (62)$$

where k_p^* and k_{ps}^* denote effective wave numbers that involve effective parameters H^* , N^* , and effective densities ρ^* , q^* . These effective parameters can be computed according to the so-called poroelastic Backus averaging.^{17,18,28} By neglecting terms higher than $O(\epsilon^2)$ —which is the overall precision of the weak-fluctuation approximation—we can replace k_{ps}^* with k_{ps} . The effective parameter H^* is computed according to

$$H^* = P_d + \alpha^2 M^* = P_d + \alpha^2 \left\langle \frac{1}{M} \right\rangle^{-1} \approx H \left[1 - \frac{\alpha^2 M}{H} B_{MM}(0) \right]. \quad (63)$$

Then the effective P wave number k_p^* can be expressed as

$$k_p^* = \omega \sqrt{\frac{\rho^*}{H^*}} \approx k_p \left[1 + \frac{1}{2} \frac{\alpha^2 M}{H} B_{MM}(0) \right]. \quad (64)$$

Using (64) Eq. (62) can be written in the form

$$\begin{aligned} \bar{k}_p^{1D} = k_p & \left[1 + \frac{\alpha^2 M}{2H} B_{MM}(0) \right. \\ & \left. + \frac{1}{2} i \frac{\alpha^2 M}{P_d} k_{ps} \int_0^\infty dz B_{MM}(z) \exp[ik_{ps}z] \right]. \quad (65) \end{aligned}$$

We will now show that the 1-D result (65) can be recovered from the 3-D result. To do so, we analyze the expression for the Q_{ik} tensor component in the wave number domain q_{33} that is of the form [cf. Eq. (B5)]

$$q_{33} \propto \int d^3 \mathbf{r} B(\mathbf{r}) \frac{\exp[ik_{ps}R]}{R} \exp[-i\mathbf{K} \cdot \mathbf{r}], \quad (66)$$

where we used only the poroelastic part of $G_{ij,ij}^F$ and $R = |\mathbf{r}|$. Note that in order to obtain Eq. (66) it is not necessary to assume statistical isotropy. We are now considering the following limiting situation.

(a) To degenerate the 3-D random medium into a 1-D random medium we stretch the correlation lengths perpendicular to the direction of wave propagation, a_\perp , to infinity so that the correlation function becomes only a function of z with parameter a_\parallel , i.e., the correlation length parallel to the direction of wave propagation. Obviously, if the wave propagates mainly in the z direction we can also write the spatial wave vector as $\mathbf{K} \approx (0, 0, k_p)^T$.

(b) Since in such a 1-D random medium there are only two directions of wave propagation ($\pm z$), we can use the small-angle approximation¹⁶ (or Fresnel approximation) of the propagator-like term $\exp[ik_{ps}R]/R$:

$$\frac{\exp[ik_{ps}R]}{R} \approx \frac{\exp[ik_{ps}z]}{z} \exp\left[\frac{ik_{ps}r_t^2}{2z}\right], \quad (67)$$

where r_t denotes the absolute value of the transverse coordinate vector $\mathbf{r}_t = (x, y)^T$. Introducing the simplifications proposed in (a) and (b) into Eq. (66) and using cylindrical coordinates $d^3 \mathbf{r} = dz dr_t d\phi r_t$, we obtain

$$\begin{aligned} q_{33} \propto 4\pi \int_0^\infty dz B(z) \frac{\exp[ik_{ps}z]}{z} \exp[-ik_p z] \\ \times \int_0^\infty dr_t r_t \exp\left[\frac{ik_{ps}r_t^2}{2z}\right]. \quad (68) \end{aligned}$$

The low-frequency condition $k_p/k_{ps} \ll 1$ means that we can replace the exponential $\exp[-ik_p z]$ by 1. The integral with respect to r always converges because k_{ps} is complex and produces an exponential decreasing multiplier. After performing the integration and substituting the result into Eq. (36), we obtain

$$\begin{aligned} \bar{k}_p = k_p & \left[1 + \frac{\alpha^2 M}{2H} B_{MM}(0) \right. \\ & \left. + \frac{1}{2} i \frac{\alpha^2 M}{P_d} k_{ps} \int_0^\infty dz B_{MM}(z) \exp[ik_{ps}z] \right], \quad (69) \end{aligned}$$

which is identical to Eq. (65).

In conclusion, the 1-D effective P wave number can be exactly obtained from the 3-D result. We note that we have only considered the particular case of M fluctuations. However, the approach can be also applied in the case of other parametrizations.

VI. CONCLUSIONS

In the framework of the theory of wave propagation in random media we analyzed the properties of the coherent wave propagating in poroelastic random media. Neglecting the ordinary elastic scattering, we only accounted for conversion scattering from fast P into Biot's slow P wave. This process of conversion scattering is equivalent to the mechanism of pore-pressure relaxation due to wave-induced perturbations. Thus, our results describe the relationship between the dynamic properties of the coherent wavefield and the mechanism of wave-induced fluid flow. In particular, we have derived an explicit expression for the effective P wave number [Eq. (49)] by applying first-order statistical smoothing of Biot's equations of poroelasticity with randomly varying coefficients. This wave number is complex and involves an integral over the correlation properties of the medium fluctuations. From this result it can be seen that the associated phase velocity is smaller than in the homogeneous background medium and that the wave is exponentially damped. We have shown that the previously reported effective P wave number for randomly layered media¹⁷ can be derived from the more general 3-D result.

Our approach is limited to the case of weak-contrast, mesoscopic inhomogeneities [conditions (58) and (61)]. In this paper we focused the analysis to wave propagation in statistically isotropic random media. However, the results can be probably generalized to the case of statistically anisotropic random media. An advantage of the statistical approach is its flexibility to handle complex geometrical distributions of the inhomogeneities. Only the spatial correlation of the fluctuations need to be known in order to compute the dynamic wavefield attributes. A detailed analysis of frequency dependencies of attenuation and dispersion due to wave-induced fluid flow and the potential applicability of the results to real rocks will be the subject of a companion paper.²⁴

ACKNOWLEDGMENTS

This work was kindly supported by the Deutsche Forschungsgemeinschaft (Contract MU 1725/1-1), CSIRO Petroleum, and the Center of Excellence for Exploration and Production Geophysics.

APPENDIX A: POROELASTIC GREEN'S TENSORS

The complete set of Green's tensors for a homogeneous and isotropic poroelastic continuum—including electroseismic coupling—was derived by Pride and Haartsen.²⁵ We reproduce only those parts of the Green's tensors, which are related to poroelastic wave propagation. Furthermore, we can simplify these tensors for low frequencies with $|k_p|/|k_{ps}| \ll 1$. We obtain

$$G_{ij}^F(\mathbf{r}-\mathbf{r}_0) = \frac{1}{4\pi\rho\omega^2} \left([k_s^2\delta_{ij} + \partial_i\partial_j] \frac{e^{ik_sR}}{R} - \partial_i\partial_j \frac{e^{ik_pR}}{R} \right) - \frac{C^2}{H^2} \frac{1}{4\pi q\omega^2} \partial_i\partial_j \frac{e^{ik_{ps}R}}{R}, \quad (\text{A1})$$

$$G_{ij}^f(\mathbf{r}-\mathbf{r}_0) = \frac{C}{H} \frac{1}{4\pi q\omega^2} \partial_i\partial_j \frac{e^{ik_{ps}R}}{R} \quad (\text{A2})$$

$$G_{ij}^w(\mathbf{r}-\mathbf{r}_0) = -\frac{1}{4\pi q\omega^2} \partial_i\partial_j \frac{e^{ik_{ps}R}}{R}, \quad (\text{A3})$$

where $R = |\mathbf{r}-\mathbf{r}_0|$. In homogeneous and isotropic media the Green's tensors only depend on R . In the low-frequency version of Biot's equations, the wave numbers of fast P , S , and slow P waves are defined as

$$k_p = \omega \sqrt{\frac{\rho}{H}}, \quad k_s = \omega \sqrt{\frac{\rho}{G}}, \quad k_{ps} = \sqrt{\frac{i\omega\eta}{\kappa_0 N}} = \omega \sqrt{\frac{q}{N}}, \quad (\text{A4})$$

where $N = MP_d/H$. Note that the first three terms of G_{ij}^F are formally identical to the elastodynamic Green's tensor.²⁹ Indeed, in the elastic limit ($K_d \rightarrow K_g$, $\alpha \rightarrow 0$ and $\phi \rightarrow 0$, $\kappa_0 \rightarrow 0$), the set of Green's tensors (A1)–(A3) reduces to the single elastodynamic Green's tensor,

$$G_{ij}^{\text{elast}}(\mathbf{r}-\mathbf{r}_0) = \frac{1}{4\pi\rho\omega^2} \left([k_s^2\delta_{ij} + \partial_i\partial_j] \frac{e^{ik_sR}}{R} - \partial_i\partial_j \frac{e^{ik_pR}}{R} \right), \quad (\text{A5})$$

where the P wave number is now given by $k_p = \omega \sqrt{\rho/(K_d + 4/3G)}$.

We define the spatial Fourier transform pair in the following way:

$$G_{ij}(\mathbf{r}-\mathbf{r}') = \int d^3\mathbf{K} g_{ij}(\mathbf{K}) \exp[i\mathbf{K} \cdot (\mathbf{r}-\mathbf{r}')], \quad (\text{A6})$$

$$g_{ij}(\mathbf{K}) = \frac{1}{(2\pi)^3} \int d^3(\mathbf{r}-\mathbf{r}') G_{ij}(\mathbf{r}-\mathbf{r}') \exp[-i\mathbf{K} \cdot (\mathbf{r}-\mathbf{r}')]. \quad (\text{A7})$$

In the wave number domain the Green's tensors (A1)–(A3) read as

$$g_{ij}^F(\mathbf{K}) = -\frac{1}{8\pi^3} \frac{1}{\rho\omega^2} \left(\frac{k_s^2\delta_{ij} - K_iK_j}{k_s^2 - K^2} + \frac{K_iK_j}{k_p^2 - K^2} \right) - \frac{1}{8\pi^3} \frac{C^2}{H^2} \frac{1}{q\omega^2} \frac{K_iK_j}{k_{ps}^2 - K^2}, \quad (\text{A8})$$

$$g_{ij}^f(\mathbf{K}) = \frac{1}{8\pi^3} \frac{C}{H} \frac{1}{q\omega^2} \frac{K_iK_j}{k_{ps}^2 - K^2}, \quad (\text{A9})$$

$$g_{ij}^w(\mathbf{K}) = -\frac{1}{8\pi^3} \frac{1}{q\omega^2} \frac{K_iK_j}{k_{ps}^2 - K^2}. \quad (\text{A10})$$

APPENDIX B: COMPUTATION OF THE OPERATOR $\mathcal{Q}_{ik}^{(1)}$

In this appendix we compute the kernel-of-mass operator $\mathcal{Q}_{ik}^{(1)}(\mathbf{r}'-\mathbf{r}'')$ defined in Eq. (37). First we note that the three terms in Eq. (37) are of the same structure,

$$\mathcal{Q}_{ik}(\mathbf{r}'-\mathbf{r}'') = \langle (L_{ij}^{(X)})' G_{jl}^Z(\mathbf{r}'-\mathbf{r}'') (L_{lk}^{(Y)})'' \rangle, \quad (\text{B1})$$

where $X=1,2$, $Y=1,2$, $Z=F,f,w$ while single and double primes indicate whether in the differential operator L_{ij} differentiation is with respect to \mathbf{r}' or \mathbf{r}'' . It is therefore sufficient to present the detailed manipulations only for one term. For example, for the differential operators $(L_{ij}^{(1)})' = \partial_i' \tilde{H}' \partial_j'$ and $(L_{ij}^{(2)})'' = \partial_i'' \tilde{C}'' \partial_j''$, we have

$$\mathcal{Q}_{ik} = \langle \partial_i' \tilde{H}' \partial_j' G_{jl}^Z(\mathbf{r}'-\mathbf{r}'') \partial_l'' \tilde{C}'' \partial_k'' \rangle. \quad (\text{B2})$$

We now use partial integrations to shift the spatial derivatives ∂_j' and ∂_l'' to the Green's function [and make use of the property $G^Z(R \rightarrow \infty) = 0$]. Averaging Eq. (B2) yields

$$\mathcal{Q}_{ik} = \partial_i' (B_{HC}(\mathbf{r}'-\mathbf{r}'') G_{jl,jl}^Z(\mathbf{r}'-\mathbf{r}'')) \partial_k'', \quad (\text{B3})$$

where we made use of the assumption of statistically homogeneous random fields [Eq. (19)]. The Fourier transform of \mathcal{Q}_{ik} is defined as

$$q_{ik}(\mathbf{K}) = \frac{1}{8\pi^3} \int d^3(\mathbf{r}'-\mathbf{r}'') \mathcal{Q}_{ik}(\mathbf{r}'-\mathbf{r}'') \exp[-i\mathbf{K} \cdot (\mathbf{r}'-\mathbf{r}'')]. \quad (\text{B4})$$

Noting that spatial derivatives correspond to multiplications in the wave number domain ($\partial_j \leftrightarrow iK_j$), we can express q_{ik} as

$$q_{ik}(\mathbf{K}) = \frac{-K_i K_k}{8\pi^3} \int d^3\mathbf{r} B_{HC}(\mathbf{r}) G_{jl,jl}^Z(\mathbf{r}) \exp[-i\mathbf{K} \cdot \mathbf{r}]. \quad (\text{B5})$$

Next, we need to compute the spatial derivatives of the Green's tensors. The poroelastic part of Green's tensor in Eq. (A1) as well as Green's tensors in Eqs. (A2) and (A3) are of the form $G_{ij}^{\text{poro}} \propto \partial_i \partial_j (e^{ik_{ps}R}/R)$. After differentiation we obtain

$$G_{ij,ij}^{\text{poro}}(\mathbf{r}) \propto -k_{ps}^2 \left(k_{ps}^2 \frac{e^{ik_{ps}R}}{R} + 4\pi\delta(\mathbf{r}) \right). \quad (\text{B6})$$

Analogously, for the derivatives of the elastic part of Green's tensor (A1) given by equation (A5) we find

$$G_{ij,ij}^{\text{elast}}(\mathbf{r}) \propto -k_p^2 \left(k_p^2 \frac{e^{ik_pR}}{R} + 4\pi\delta(\mathbf{r}) \right). \quad (\text{B7})$$

Note that the derivatives in (B6) and (B7) are of the same functional form, despite the fact that G_{ij}^{elast} includes S wave contributions. These contributions, however, cancel out when $G_{ij,ij}^{\text{elast}}$ is computed. The parts of the differential operator $L_{ik}^{(1)}$ involving fluctuations of the shear modulus require the computation of the terms $G_{33,33}^{\text{elast}}$ and $G_{3j,3j}^{\text{elast}}$. It can be shown that these derivatives can be expressed in terms of $G_{ij,ij}$:

$$G_{33,33} = G_{ij,ij} \left(\frac{1}{2} \int_0^\pi d\theta \sin\theta \cos^4\theta + O(k_p^2 R^2) \right). \quad (\text{B8})$$

$$= G_{ij,ij} \left(\frac{1}{5} + O(k_p^2 R^2) \right). \quad (\text{B9})$$

Analogously,

$$G_{3j,3j} = G_{ij,ij} \left(\frac{1}{3} + O(k_p^2 R^2) \right). \quad (\text{B10})$$

Neglecting terms of the order $O(k_p^2 R^2)$ will produce no additional inaccuracy because of the restriction to low frequencies. Because of the same argument, we also neglect the first term in (B7). It is this term that is responsible for dynamic effects of elastic scattering and that is left out in the present analysis.

Inserting expressions (B6) and (B7) into Eq. (B5), introducing spherical coordinates, assuming statistical isotropy [$B(\mathbf{r})=B(|\mathbf{r}|)$], and integrating over the angular coordinates, we obtain

$$q_{33} = -\frac{1}{4\pi^3} \frac{C^2}{N} \left(k_p^2 B_{HC}(0) + k_p k_{ps}^2 \int_0^\infty dr B_{HC}(r) \exp[ik_{ps}r] \sin(k_p r) \right). \quad (\text{B11})$$

In accordance with our low-frequency assumption, we replace the sine function by its argument and obtain Eq. (45). Repeating these computations for the remaining terms involving correlations and cross-correlation between the random fields of H, G , and C yields Eqs. (43)–(48).

¹G. Mavko and A. Nur, “Melt squirt in aethenosphere,” *J. Geophys. Res.* **80**, 1444–1448 (1975).
²W. F. I. Murphy, K. W. Winkler, and R. L. Kleinberg, “Acoustic relaxation in sedimentary rocks: Dependence on grain contacts and fluid saturation,” *Geophysics* **51**, 757–766 (1986).
³T. Jones, “Pore fluids and frequency dependent wave propagation in rocks,” *Geophysics* **51**, 1939–1953 (1986).
⁴J. E. White, *Underground Sound: Application of Seismic Waves* (Elsevier, Amsterdam, 1983).
⁵W. F. Murphy, “Effects of partial water saturation on attenuation in Massillon sandstone and Vycor porous glass,” *J. Acoust. Soc. Am.* **71**, 1458–1468 (1982).
⁶G. A. Gist, “Interpreting laboratory velocity measurements in partially gas-saturated rocks,” *Geophysics* **59**, 1100–1109 (1994).
⁷R. Knight, J. Dvorkin, and A. Nur, “Acoustic signatures of partial saturation,” *Geophysics* **63**, 132–138 (1998).
⁸M. A. Biot, “Mechanics of deformation and acoustic propagation in porous media,” *J. Appl. Phys.* **33**, 1482–1498 (1962).
⁹J. E. White, “Computed seismic speeds and attenuation in rocks with partial gas saturation,” *Geophysics* **40**, 224–232 (1975).
¹⁰J. E. White, N. G. Mikhaylova, and F. M. Lyakhovitsky, “Low-frequency seismic waves in fluid saturated layered rocks,” *Izv., Acad. Sci., USSR, Phys. Solid Earth* **11**, 654–659 (1975).

¹¹N. C. Dutta and H. Ode, “Attenuation and dispersion of compressional waves in fluid-filled porous rocks with partial gas saturation (White model)—Part I: Biot theory,” *Geophysics* **44**, 1777–1788 (1979).
¹²S. R. Pride and J. G. Berryman, “Linear dynamics of double-porosity dual-permeability materials I. governing equations and acoustic attenuation,” *Phys. Rev. E* **68**, 036603 (2003).
¹³S. R. Pride, J. G. Berryman, and J. M. Harris, “Seismic attenuation due to wave-induced flow,” *J. Geophys. Res.* **109**, B01201 (2004).
¹⁴D. L. Johnson, “Theory of frequency dependent acoustics in patchy-saturated porous media,” *J. Acoust. Soc. Am.* **110**, 682–694 (2001).
¹⁵S. L. Lopatnikov and B. Gurevich, “Transformational mechanism of elastic wave attenuation in saturated porous media,” *Izv., Acad. Sci., USSR, Phys. Solid Earth* **24**, 151–154 (1988).
¹⁶S. M. Rytov, Y. A. Kravtsov, and V. I. Tatarskii, “Wave propagation through random media,” Vol. 4 of *Principles of Statistical Radiophysics* (Springer-Verlag, Berlin, 1989).
¹⁷B. Gurevich and S. L. Lopatnikov, “Velocity and attenuation of elastic waves in finely layered porous rocks,” *Geophys. J. Int.* **121**, 933–947 (1995).
¹⁸A. N. Norris, “Low-frequency dispersion and attenuation in partially saturated rocks,” *J. Acoust. Soc. Am.* **94**, 359–370 (1993).
¹⁹B. Gurevich, A. P. Sadovnichaja, S. L. Lopatnikov, and S. A. Shapiro, “Scattering of a compressional wave in a poroelastic medium by an ellipsoidal inclusion,” *Geophys. J. Int.* **133**, 91–103 (1998).
²⁰J. B. Keller, “Stochastic equations and wave propagation in random media,” *Proc. Symp. Appl. Math.* **16**, 145–170 (1964).
²¹F. C. Karal and J. B. Keller, “Elastic, electromagnetic and other waves in random media,” *J. Math. Phys.* **5**, 537–547 (1964).
²²J. J. McCoy, “On the dynamic response of disordered composites,” *J. Appl. Mech.* **40**, 511–517 (1973).
²³N. Gold, S. A. Shapiro, S. Bojinski, and T. M. Müller, “An approach to upscaling for seismic waves in statistically isotropic heterogeneous elastic media,” *Geophysics* **65**, 1837–1850 (2000).
²⁴T. M. Müller and B. Gurevich, “Wave-induced fluid flow in porous random media: Attenuation and dispersion of seismic waves,” *J. Acoust. Soc. Am.*, accepted.
²⁵S. R. Pride and M. W. Haartsen, “Electroseismic wave properties,” *J. Acoust. Soc. Am.* **100**, 1301–1315 (1996).
²⁶A. N. Norris, “Radiation from a point source and scattering theory in a fluid-saturated porous solid,” *J. Acoust. Soc. Am.* **77**, 2012–2023 (1985).
²⁷B. Gurevich and S. L. Lopatnikov, “Attenuation of longitudinal waves in a saturated porous medium with random inhomogeneities,” *Transactions (Doklady) of the USSR Academy of Sciences Earth Sci. Sections* **281**, 47–50 (1985).
²⁸S. Gelinsky and S. A. Shapiro, “Poroelastic Backus-averaging for anisotropic, layered fluid and gas saturated sediments,” *Geophysics* **62**, 1867–1878 (1997).
²⁹J. E. Gubernatis, E. Domany, and J. A. Krumhansl, “Formal aspects of the theory of the scattering of ultrasound by flaws in elastic materials,” *J. Appl. Phys.* **48**, 2804–2811 (1977).

Linear and nonlinear propagation of higher order modes in hard-walled circular ducts containing a real gas

Stefan Scheichl^{a)}

Acoustics Research Institute, Austrian Academy of Sciences, A-1010 Vienna, Austria

(Received 10 September 2004; revised 14 January 2005; accepted 21 January 2005)

This paper deals with the linear and nonlinear propagation of sound waves through a real gas contained in a circular tube with rigid, isothermal walls. Special emphasis is placed on the asymptotically correct treatment of the higher order modes and their interaction with the acoustic boundary layer. In the first part, a linear perturbation analysis is carried out to calculate the correction terms arising from the viscothermal damping mechanisms present in the system. In extension to previous work, the propagation length is assumed to be so large that the exponentially growing boundary layer effects do not only affect the second order terms of the sound pressure but also the leading order terms. The series expansions derived for the propagation parameters extend the results given in the literature with additional terms resulting from viscosity and heat conduction in the core region. The second part is concerned with the nonlinear modulation of a wave packet transmitted through a real gas. A damped nonlinear Schrödinger equation is derived and its solutions for positive as well as negative values of the nonlinearity parameter are studied. In particular, the case of wave propagation in ducts containing a so-called BZT fluid is discussed. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1869113]

PACS numbers: 43.25.Cb, 43.25.Ba, 43.20.Mv, 43.20.Hq [MO]

Pages: 1806–1827

I. INTRODUCTION

Since the end of the 19th century the problem of extending the classical Kirchhoff “wide tube” result¹ for the viscothermal attenuation of plane waves, which was derived from a general dispersion relation, to other cases, such as narrow tubes, higher order modes, and finite amplitude sound waves propagation, has been subject of a number of papers. Rayleigh² was one of the first to obtain useful approximations for the narrow tube, i.e., the case where the driving frequencies are so low that the viscothermal boundary extends itself over the whole tube diameter. In 1949 Zwicker and Kosten³ introduced an approximate theory based on the so-called low reduced frequency assumptions that enabled the simplification of the basic equations such that the transmission line parameters of the fundamental mode could be given in closed form for a wide range of driving frequencies. A thorough discussion of this approach including comparisons with the investigations of other authors and with numerical solutions of Kirchhoff’s general dispersion equations can be found in the 1975 paper by Tijdeman (see Ref. 4 and the references therein). Later Keefe⁵ derived more accurate higher order series expansions of the plane wave transmission line parameters for the cases of high and low frequency waves. The applicability and validity of these results from the viewpoint of a systematic perturbation analysis have recently been studied by Scheichl.⁶

By applying the method of boundary layer admittance, Beatty⁷ calculated the leading order boundary layer attenuation caused by the viscous and thermal losses of all higher modes which opened the way to more refined analyses with respect to higher order corrections and the frequency range

around the cutoff frequencies (see, e.g., Refs. 8 and 9). A detailed investigation on that topic within the framework of a linear multiple scales perturbation theory is presented in Ref. 6. However, this study is confined to the case of axisymmetric wave motion.

Moreover, in many physical problems, the sound pressure levels involved might become so high that the problem cannot be treated using a linear analysis. Again, by adopting the assumption that the frequencies are sufficiently high such that the flow field can be split into a thin viscous wall layer and an essentially inviscid core layer, Chester¹⁰ and independently also Blackstock¹¹ were able to derive a generalized Burgers equation for (almost) plane waves where the boundary layer effects are described by a dissipative-dispersive memory term. It should be mentioned that the same equation may also be used to calculate the slowly time-varying distortions of the profiles of counterpropagating waves and the buildup of standing nonlinear oscillations in resonators (cf. the articles by Gusev *et al.*¹² and Menguy and Gilbert¹³). The experiments reported in the latter paper are in good accordance with the theoretically predicted pressure amplitudes. In extension to Blackstock’s studies, by adopting the Khokhlov-Zabolotskaya-Kuznetsov model, Makarov and Vatrushina¹⁴ analyzed the case of the wavelengths being so short that the losses in the core region are of the same order of magnitude as the boundary layer effects and, after averaging the sound pressure over the cross section, arrived at a modified Blackstock equation. For a more in-depth discussion of these and other theoretical models, the reader is referred to the review paper by Makarov and Ochmann.¹⁵

As far as the nonlinear propagation of higher order modes in circular ducts is concerned, the investigation by Keller and Millman¹⁶ should be mentioned, who, using the method of strained parameters, determined the nonlinear

^{a)}Electronic mail: stefan.scheichl@oew.ac.at

wave number shift for the symmetrical modes. In the 1998 paper by Foda¹⁷ the same technique is used in order to investigate the nonlinear interaction of different modes. Furthermore, Nayfeh¹⁸ derived a nonlinear Schrödinger equation for the propagation of a wave packet centered at the wave number and the frequency of a higher order carrier mode by applying the method of multiple scales. This is the study the second part of this present paper is based on.

In contrast to previous work where the fluid, in general, is assumed to be perfect gas, the investigations presented here are concerned with the physical processes taking place when sound is transmitted through a real gas. Both, linear as well as nonlinear wave propagation in a hard walled circular tube, will be studied. In order to ensure that the acoustic motion is not dominated by thermoviscous effects, the driving frequencies are presupposed to be sufficiently high such that the spatial range consumed by the boundary layer is small compared to the tube diameter. Consistent with that, the characteristic wavelength is assumed to be of the same order of magnitude as the diameter or smaller and, consequently, the occurrence of higher order modes has to be taken into account. The paper is structured as follows.

The first part deals with the linear wave propagation under the action of boundary layer and core layer attenuation. The primary aim pursued in this study is to derive the asymptotically correct solutions for the involved field quantities *including* the second-order terms that are uniformly valid over a considerably larger spatial range than that constituted by the wavelengths. To this end, the *method of multiple scales* is applied. Additionally, the *method of matched asymptotic expansions* is used, in order to analyze the changes in lateral direction within the boundary layer. As already shown in Ref. 6 for the symmetrical modes, a higher-order perturbation analysis also incorporating the dissipative losses in the core region turns out to be necessary if the length of the tube becomes so large that the exponentially growing effects due to viscosity and heat conduction in the wall layer do not only affect the second-order correction terms of the sound pressure but also the leading order terms. The thus obtained results extend the solutions for the radial modes given in Ref. 6 with the corresponding expressions for the asymmetrical modes and, furthermore, also account for the effects arising from the adoption of a general equation of state.

In the second part, beginning with Sec. IV, the same perturbation techniques as before are used to derive an envelope evolution equation for the temporal and spatial modulation of nonlinear, dispersive waves propagating in a long circular duct containing a real gas. This equation will turn out to be a nonlinear Schrödinger equation, extended with a dissipative damping term resulting from the thermoviscous boundary layer effects. In addition, the expressions for the second-order correction terms are calculated.

As mentioned before, a similar analysis, however, restricted to the case of a perfect gas was carried out by Nayfeh.¹⁸ Extending this study to real gases is motivated by the following observations: It is well known that the propagation of weakly nonlinear plane waves in homogeneous, nondissipative media can be described by an inviscid Bur-

gers equation in which the so-called *fundamental derivative* (at constant entropy \bar{s}),

$$\Gamma = \frac{1}{\bar{c}} \left(\frac{\partial(\bar{\rho}\bar{c})}{\partial\bar{\rho}} \right)_{\bar{s}} = 1 + \frac{B}{2A}, \quad (1)$$

accounts for the rate of change of the sound speed \bar{c} with density $\bar{\rho}$ (see, e.g., Ref. 15). In this definition, as usual, the ratio B/A denotes the so-called *acoustic nonlinearity parameter*, whereas the quantity Γ is sometimes also referred to as the *intrinsic gasdynamic nonlinearity parameter*. For perfect gases with constant specific heats, $\Gamma = (\gamma + 1)/2 > 0$, where γ represents the ratio of the specific heats, and the fluid motion is said to exhibit positive nonlinearity. In this case only *compression shocks* can form and propagate in the gas. The possibility that also *rarefaction shocks* can form in real fluids, i.e., that Γ might become negative, seems to have been recognized first by Bethe¹⁹ and independently by Zel'dovich.²⁰ A class of fluids involving gases with high specific heats can be identified for which the curvature of the isentropes in the pressure-density state space is reversed near the coexistence curve in the vicinity of the critical point and, consequently, Γ changes sign. More sophisticated studies based on the Martin-Hou equation of state are due to Thompson and co-workers,^{21,22} who gave specific examples of such fluids, which include hydrocarbons and fluorocarbons of moderate complexity. In recognition of these investigations, fluids having the distinguishing property that the fundamental derivative can change sign are commonly referred to as BZT fluids. Two impressive analyses of several commercially available fluorocarbons by means of the Martin-Hou equation are given in the 1989 paper by Cramer²³ and the 1997 paper by Monaco *et al.*²⁴ and clearly substantiate the existence of negative nonlinearity regions. Recently, also fluids belonging to the group of the so-called siloxanes were shown to be possible candidates for media with a negative Γ region in the gaseous state.²⁵ Furthermore, Kluwick²⁶ demonstrated that if the equilibrium state of the BZT fluid is chosen to be in the neighborhood to the transition line where $\Gamma = 0$, the propagation of planar, weakly nonlinear acoustic waves in rigid tubes is governed by a modified Burgers equation very similar to that derived by Blackstock,¹¹ however, extended with an additional cubic nonlinearity term.

The consequence of the above discussion is that the possibility of $\Gamma < 0$ obviously provides new physical phenomena in gas dynamics and acoustics and, thus, may also be of importance for the case considered in the second part, namely, the evolution of a wave packet centered at the wave number and the frequency of a higher order carrier mode. As it will turn out, the type of nonlinearity is then not only determined by the fundamental derivative, but also by its first derivative with respect to the density.

II. PROBLEM FORMULATION

In this paper, the linear and the weakly nonlinear wave propagation in a rigid circular tube with isothermal walls containing a real gas will be investigated. The problem considered and the notation used are shown in Fig. 1.

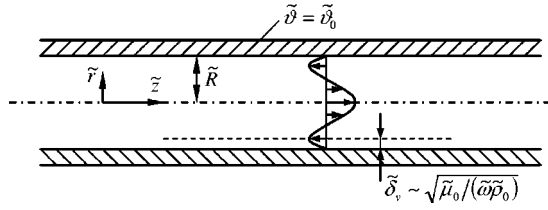


FIG. 1. Sketch showing duct geometry and notation.

A. General assumptions

Throughout the investigations, the following assumptions are adopted:

$$\text{Re} = \frac{\bar{c}_0 \bar{\lambda} \bar{\rho}_0}{\bar{\mu}_0} \gg 1, \quad l = \frac{\bar{R}}{\bar{\lambda}} \gg \frac{1}{\sqrt{\text{Re}}}, \quad M = \frac{\bar{W}}{\bar{c}_0} \ll 1. \quad (2)$$

Here, the quantities $\bar{c} = \sqrt{(\partial \bar{p} / \partial \bar{\rho})_{\bar{s}}}$, \bar{p} , $\bar{\rho}$, \bar{s} , $\bar{\mu}$, $\bar{\lambda}$, \bar{R} , and \bar{W} , and the parameters Re , l , and M denote the speed of sound, the fluid pressure, the density of the fluid, the specific entropy, the dynamic viscosity, the characteristic wavelength, the radius of the tube, the characteristic particle velocity, the acoustic Reynolds number, the reduced frequency,⁴ and the Mach number. Quantities evaluated at the equilibrium reference state are indicated by the subscript 0.

Furthermore, it is assumed that the fluid is a gas such that the equilibrium quantities satisfy

$$\begin{aligned} \bar{\beta}_0 \bar{\vartheta}_0 &= O(1), \quad \gamma_0 - 1 = \frac{\bar{C}_{p0}}{\bar{C}_{v0}} - 1 = O(1), \\ \text{Pr} &= \frac{\bar{\mu}_0 \bar{C}_{p0}}{\bar{\kappa}_0} = O(1), \end{aligned} \quad (3)$$

$$\Gamma_0 = O(1), \quad \Lambda_0 = \left[\bar{\rho} \left(\frac{\partial \Gamma}{\partial \bar{\rho}} \right)_{\bar{s}} \right]_0 = 0(1),$$

where $\bar{\beta} = -(1/\bar{\rho})(\partial \bar{p} / \partial \bar{\vartheta})_{\bar{p}}$, $\bar{\vartheta}$, γ , $\bar{\kappa}$, Pr , and Λ represent the coefficient of thermal expansion, the temperature, the ratio of the specific heats, the thermal conductivity, the Prandtl number, and the first derivative of the fundamental derivative (1).

Since $\text{Pr} = O(1)$, the thermal and the viscous boundary layer thicknesses are of the same order of magnitude (see, e.g., Ref. 15):

$$\bar{\delta}_r \sim \bar{\delta}_v \sim \sqrt{\frac{\bar{\mu}_0}{\bar{\omega} \bar{\rho}_0}} \sim \frac{\bar{\lambda}}{\sqrt{\text{Re}}}, \quad (4)$$

with $\bar{\omega}$ being the radian driving frequency. Consequently, the first and second restrictions in Eq. (2) can be interpreted as defining two ordering relationships for the most relevant geometrical scales, i.e., the wavelength, the radius, and the spatial range consumed by the boundary layer: As a result of

$$\text{Re} \sim \frac{\bar{\lambda}^2}{\bar{\delta}_v^2} \gg 1, \quad l^2 \text{Re} \sim \text{He}^2 \text{Re} \sim \text{St}^2 = \frac{\bar{R}^2}{\bar{\delta}_v^2} \gg 1, \quad (5)$$

where

$$\text{He} = \frac{\bar{\omega} \bar{R}}{\bar{c}_0}, \quad \text{St} = \sqrt{\frac{\bar{\omega} \bar{\rho}_0 \bar{R}^2}{\bar{\mu}_0}}, \quad (6)$$

the quantity $\bar{\delta}_v$ turns out to be small compared to the characteristic wavelength as well as the radius. Here, the quantities $\text{He} \sim l$ and St are the Helmholtz number and the Stokes number, respectively.

In order to capture the most relevant effects arising from the excitation of higher order radial and azimuthal modes, the linear investigations in Sec. III will proceed from the assumption that the characteristic wavelength $\bar{\lambda}$ is comparable to the diameter of the duct, i.e., $l = O(1)$, whereas the propagation length will be presupposed to be much larger than $\bar{\lambda}$. In extension to that, the case of \bar{R} being much larger than the wavelengths, will be considered in Sec. III B. As mentioned in the Introduction, the opposite case, i.e., linear plane wave propagation with $l \ll 1$, has already been extensively discussed in the literature.

For the study in Sec. IV dealing with the nonlinear propagation of a wave packet, it will again be assumed that $\bar{R} \sim \bar{\lambda}$. However, as explicated in Ref. 18, the fundamental mode has to be excluded as a potential carrier mode. This is due to the fact that it propagates as a nearly plane wave and, thus, at the same speed as its harmonics. Consequently, it cannot be excited without strongly exciting the higher harmonics as well since the effect of gas nonlinearity leads to a cumulative distortion of the wave and a shift of energy among the several harmonics. Analyzing this case within the framework of an asymptotic perturbation theory would require the application of the plane wave evolution equation proposed by Blackstock¹¹ and others^{10,12,13} or, respectively, its generalized version by Kluwick²⁶ (see the Introduction).

B. Basic equations and perturbation approach

It is reasonable to formulate the problems under consideration in terms of nondimensional quantities by involving the wavelength $\bar{\lambda}$, the radius \bar{R} , and the equilibrium quantities $\bar{\rho}_0$, \bar{c}_0 , $\bar{\beta}_0$, \bar{C}_{p0} , and $\bar{\mu}_0$ introduced above:

$$\begin{aligned} z &= \frac{\bar{z}}{\bar{\lambda}}, \quad r = \frac{\bar{r}}{\bar{R}}, \quad t = \frac{\bar{t} \bar{c}_0}{\bar{\lambda}}, \quad \omega = \frac{\bar{\omega} \bar{\lambda}}{\bar{c}_0}, \quad v_z = \frac{\bar{v}_z}{\bar{c}_0}, \\ v_r &= \frac{\bar{v}_r}{\bar{c}_0}, \quad v_\phi = \frac{\bar{v}_\phi}{\bar{c}_0}, \quad p = \frac{\bar{p}}{\bar{\rho}_0 \bar{c}_0^2}, \\ \rho &= \frac{\bar{\rho}}{\bar{\rho}_0}, \quad \vartheta = \bar{\vartheta} \bar{\beta}_0, \quad c = \frac{\bar{c}}{\bar{c}_0}, \quad s = \frac{\bar{s}}{\bar{C}_{p0}}, \\ \beta &= \frac{\bar{\beta}}{\bar{\beta}_0}, \quad C_p = \frac{\bar{C}_p}{\bar{C}_{p0}}, \quad \eta = \frac{\bar{\eta}}{\bar{\mu}_0}. \end{aligned} \quad (7)$$

Here, v_z , v_r , v_ϕ , and η denote, respectively, the velocities in axial, radial, and azimuthal directions, and the bulk viscosity. Furthermore, the scaling parameter

$$\alpha = d \frac{\bar{\lambda}}{\bar{R} \sqrt{\text{Re}}} = \frac{d}{l \sqrt{\text{Re}}} \quad (8)$$

is introduced where d is an arbitrary constant of order $O(1)$, which, together with Eqs. (2) and (5), leads to the relationship $\alpha \sim \text{St}^{-1}$.

In the following analysis it will be assumed that the variations of the thermal conductivity and the dynamic viscosities are so small that these quantities can be regarded as constant, i.e., $\tilde{\kappa} = \tilde{\kappa}_0$, $\tilde{\mu} = \tilde{\mu}_0$, and $\tilde{\eta} = \tilde{\eta}_0$. However, it should be emphasized that due to the assumption of a very small Mach number [see Eq. (2)] the results derived below would remain unchanged even if the commonly used approximative power laws $\tilde{\kappa} = \tilde{\kappa}_0(\tilde{\vartheta}/\tilde{\vartheta}_0)^\sigma$ and $\tilde{\mu} = \tilde{\mu}_0(\tilde{\vartheta}/\tilde{\vartheta}_0)^\sigma$, where the coefficient $\sigma = O(1)$, were adopted.

The three-dimensional Navier-Stokes equations in cylindrical coordinates, the continuity equation, the energy equation, and the general equation of state then read

$$\rho \frac{dv_z}{dt} = \frac{\alpha^2 l^2}{d^2} \left(\frac{4}{3} + \eta_0 \right) \frac{\partial^2 v_z}{\partial z^2} + \frac{\alpha^2}{d^2} \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial v_z}{\partial r} \right) + \frac{\alpha^2}{d^2} \frac{1}{r^2} \frac{\partial^2 v_z}{\partial \phi^2} + \frac{\alpha^2 l}{d^2} \left(\frac{1}{3} + \eta_0 \right) \times \frac{1}{r} \left[\frac{\partial}{\partial r} \left(r \frac{\partial v_r}{\partial z} \right) + \frac{\partial^2 v_\phi}{\partial z \partial \phi} \right] - \frac{\partial p}{\partial z}, \quad (9)$$

$$\rho \left(\frac{dv_r}{dt} - \frac{1}{l} \frac{v_\phi^2}{r} \right) = \frac{\alpha^2 l^2}{d^2} \frac{\partial^2 v_r}{\partial z^2} + \frac{\alpha^2}{d^2} \left(\frac{4}{3} + \eta_0 \right) \frac{1}{r} \left[\frac{\partial}{\partial r} \left(r \frac{\partial v_r}{\partial r} \right) - \frac{v_r}{r} \right] + \frac{\alpha^2}{d^2} \frac{1}{r^2} \frac{\partial^2 v_r}{\partial \phi^2} + \frac{\alpha^2 l}{d^2} \left(\frac{1}{3} + \eta_0 \right) \frac{\partial^2 v_z}{\partial z \partial r} + \frac{\alpha^2}{d^2} \left(\frac{1}{3} + \eta_0 \right) \times \frac{1}{r} \left(\frac{\partial^2 v_\phi}{\partial r \partial \phi} + \frac{1}{r} \frac{\partial v_\phi}{\partial \phi} \right) - \frac{\alpha^2}{d^2} \frac{2}{r^2} \frac{\partial v_\phi}{\partial \phi} - \frac{1}{l} \frac{\partial p}{\partial r}, \quad (10)$$

$$\rho \left(\frac{dv_\phi}{dt} + \frac{1}{l} \frac{v_r v_\phi}{r} \right) = \frac{\alpha^2 l^2}{d^2} \frac{\partial^2 v_\phi}{\partial z^2} + \frac{\alpha^2}{d^2} \frac{1}{r} \left[\frac{\partial}{\partial r} \left(r \frac{\partial v_\phi}{\partial r} \right) - \frac{v_\phi}{r} \right] + \frac{\alpha^2}{d^2} \left(\frac{4}{3} + \eta_0 \right) \frac{1}{r^2} \frac{\partial^2 v_\phi}{\partial \phi^2} + \frac{\alpha^2 l}{d^2} \left(\frac{1}{3} + \eta_0 \right) \frac{1}{r} \frac{\partial^2 v_z}{\partial z \partial \phi} + \frac{\alpha^2}{d^2} \left(\frac{1}{3} + \eta_0 \right) \frac{1}{r^2} \frac{\partial}{\partial r} \left(r \frac{\partial v_r}{\partial \phi} \right) + \frac{\alpha^2}{d^2} \frac{2}{r^2} \frac{\partial v_r}{\partial \phi} - \frac{1}{l r} \frac{\partial p}{\partial \phi}, \quad (11)$$

$$\frac{d\rho}{dt} + \rho \nabla \cdot \mathbf{v} = 0, \quad (12)$$

$$\rho \vartheta \frac{ds}{dt} = \rho C_p \frac{d\vartheta}{dt} - G_0 \vartheta \beta \frac{d\rho}{dt} = \frac{\alpha^2 l^2}{d^2 \text{Pr}} \Delta \vartheta + \frac{\alpha^2 l^2}{d^2} (\gamma - 1) \Phi, \quad (13)$$

$$c^2 (\delta\rho) = \left(1 + G_0 \frac{c^2 \vartheta \beta^2}{C_p} \right) (\delta p) - c^2 \rho \beta (\delta \vartheta). \quad (14)$$

The quantity Φ appearing in the last term in Eq. (13) is the so-called dissipation function nondimensionalized by $(\tilde{c}_0^2 \tilde{\mu}_0)/\tilde{\lambda}^2$; its explicit form can be found, e.g., in Ref. 27. However, the term $\alpha^2 \Phi$ will turn out to be negligibly small. Furthermore, d/dt denotes the substantial derivative of a scalar A ,

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} + \mathbf{v} \cdot \nabla A, \quad \text{with } \nabla = \left(\frac{\partial}{\partial z}, \frac{1}{l} \frac{\partial}{\partial r}, \frac{1}{lr} \frac{\partial}{\partial \phi} \right)^T, \quad (15)$$

and the divergence of a vector \mathbf{A} and the Laplacian are defined by

$$\nabla \cdot \mathbf{A} = \left(\frac{\partial A_z}{\partial z} + \frac{1}{lr} \frac{\partial}{\partial r} (r A_r) + \frac{1}{lr} \frac{\partial A_\phi}{\partial \phi} \right), \quad \Delta A = \nabla \cdot \nabla A. \quad (16)$$

Please note that in Eqs. (13) and (14) the nondimensional Grüneisen parameter G ,

$$G = \frac{\tilde{p}}{\tilde{\vartheta}} \left(\frac{\partial \tilde{\vartheta}}{\partial \tilde{p}} \right)_{\tilde{s}} = \frac{\tilde{c}^2 \tilde{\beta}}{\tilde{C}_p} = \frac{\gamma - 1}{\tilde{\vartheta} \tilde{\beta}} = \frac{\gamma - 1}{\vartheta \beta}, \quad (17)$$

evaluated at the unperturbed reference state has been introduced; Eq. (3) shows that $G_0 = O(1)$. In the case of a perfect gas, one obtains $c^2 = \gamma p / \rho$, $\beta = 1 / \vartheta$, $c^2 \vartheta \beta^2 / C_p = G / G_0$, $G = \gamma - 1$ and, as expected, Eq. (14) reduces to

$$\frac{(\delta\rho)}{\rho} = \frac{(\delta p)}{p} - \frac{(\delta\vartheta)}{\vartheta} \Rightarrow \frac{p}{\rho \vartheta} = \frac{1}{\gamma_0} = \text{const.} \quad (18)$$

Equations (9)–(14) will be solved subject to the boundary conditions at the tube wall,

$$r = 1: \quad v_z = v_r = v_\phi = 0, \quad \vartheta = \tilde{\vartheta}_0 \tilde{\beta}_0, \quad (19)$$

and the condition that the solutions are bounded at the center of the tube.

Utilizing the parameter introduced in Eq. (8) and assuming that $\tilde{\lambda} = O(\tilde{R})$, the conditions (2) and (5) can be written as

$$\alpha \ll 1, \quad l = O(1), \quad M \ll 1 \quad (20)$$

since then $\text{Re}^{-1} \sim \alpha^2 \ll 1$ and $\text{St}^{-1} \sim \alpha \ll 1$, suggesting the use of M and α as perturbation parameters for an asymptotic analysis. Hence, the velocity components and the relevant thermodynamic quantities are expressed in the form

$$\begin{aligned} v_i &= M(v_{i1} + \alpha v_{i\alpha} + \alpha^2 v_{i\alpha^2} + \dots) \\ &\quad + M^2(v_{iM} + \alpha v_{iM\alpha} + \dots) + \dots, \\ p &= \frac{\tilde{p}_0}{\tilde{\rho}_0 \tilde{c}_0^2} + M(p_1 + \alpha p_\alpha + \alpha^2 p_{\alpha^2} + \dots) \\ &\quad + M^2(p_M + \alpha p_{M\alpha} + \dots) + \dots, \\ \rho &= 1 + M(\rho_1 + \alpha \rho_\alpha + \alpha^2 \rho_{\alpha^2} + \dots) \\ &\quad + M^2(\rho_M + \alpha \rho_{M\alpha} + \dots) + \dots, \\ \vartheta &= \tilde{\vartheta}_0 \tilde{\beta}_0 + M(\vartheta_1 + \alpha \vartheta_\alpha + \alpha^2 \vartheta_{\alpha^2} + \dots) \\ &\quad + M^2(\vartheta_M + \alpha \vartheta_{M\alpha} + \dots) + \dots, \end{aligned} \quad (21)$$

$$c = 1 + Mc_1 + \dots, \quad \beta = 1 + M\beta_1 + \dots,$$

$$C_p = 1 + MC_{p1} + \dots,$$

where $i = z, r, \phi$, respectively. The sound pressure is thus given by $p_s = p - \bar{p}_0 / (\bar{\rho}_0 \bar{c}_0^2)$.

As will be shown in the following, significant simplifications of these expansions are possible if either the Mach number is assumed to be very small (linear case) or M and α can be related to each other such that only one perturbation parameter is necessary (nonlinear case).

III. THE LINEAR CASE

In order to avoid nonlinear effects entering the resulting equations or, more precisely, those that are relevant to solving the problem, the Mach number is assumed to satisfy

$$M \ll \alpha^2. \quad (22)$$

Consequently, each unsteady perturbation w can be decomposed into a steady modal amplitude and a time-harmonic function such that

$$w(z, r, \phi, t) = \Re(\hat{w}(z, r, \phi)e^{j\omega t}). \quad (23)$$

Here and in the following, the accents “ $\hat{}$ ” denote complex-valued quantities.

As shown in Ref. 6, the investigation of the most significant features of the leading order and second-order terms generated by acoustic waves emerging over spatial ranges in the longitudinal direction of the order $O(\alpha^{-1})$ requires the introduction of the length scales z , $z_1 = \alpha z$, and $z_2 = \alpha^2 z$ according to

$$\hat{w}(z, r, \phi) \rightarrow \hat{w}(z, z_1, z_2, r, \phi), \quad \frac{\partial}{\partial z} \rightarrow \frac{\partial}{\partial z} + \alpha \frac{\partial}{\partial z_1} + \alpha^2 \frac{\partial}{\partial z_2}, \quad (24)$$

because otherwise the occurrence of secular terms could not be avoided.

Substitution of the expressions (21) into the system (9)–(14) generates a set of equations which is valid in the entire width of the tube except the boundary region and, therefore, is called the *outer* expansion. Close to the tube wall, where the stretched lateral coordinate

$$y = \frac{1-r}{\alpha} \quad (25)$$

is of $O(1)$, viscosity and heat conduction play an important role; these effects have to be accounted for by a separate investigation of the boundary layer. Consequently, for the *inner* expansion, the coordinate r then has to be replaced with $1 - \alpha y$ and, furthermore,

$$\frac{\partial}{\partial r} \rightarrow -\frac{1}{\alpha} \frac{\partial}{\partial y}. \quad (26)$$

Please note that in the inner expansion, the density terms will be denoted by Ψ , whereas all other inner quantities will be written in capital letters, for example $V_{z\epsilon}$. In order to match the quantities arising from the two expansions, Van Dyke's *asymptotic matching principle* (see, e.g., Ref. 28, pp. 173–179) will be applied.

Evaluation of the continuity equation (12) together with the boundary conditions (19) then yields

$$\frac{\partial \hat{V}_{r1}}{\partial y} = 0 \Rightarrow \hat{V}_{r1} = 0, \quad (27)$$

$$j\omega \hat{\Psi}_1 + \frac{\partial \hat{V}_{z1}}{\partial z} + \frac{1}{l} \left(-\frac{\partial \hat{V}_{r\alpha}}{\partial y} + \frac{\partial \hat{V}_{\phi 1}}{\partial \phi} \right) = 0, \quad (28)$$

$$j\omega \hat{\Psi}_\alpha + \frac{\partial \hat{V}_{z\alpha}}{\partial z} + \frac{\partial \hat{V}_{z1}}{\partial z_1} + \frac{1}{l} \left(-\frac{\partial \hat{V}_{r\alpha^2}}{\partial y} + \hat{V}_{r\alpha} + \frac{\partial \hat{V}_{\phi\alpha}}{\partial \phi} + y \frac{\partial \hat{V}_{\phi 1}}{\partial \phi} \right) = 0, \quad (29)$$

$$j\omega \hat{\rho}_1 + \nabla \cdot \hat{\mathbf{v}}_1 = 0, \quad (30)$$

$$j\omega \hat{\rho}_\alpha + \frac{\partial \hat{v}_{z1}}{\partial z_1} + \nabla \cdot \hat{\mathbf{v}}_\alpha = 0, \quad (31)$$

$$j\omega \hat{\rho}_{\alpha^2} + \frac{\partial \hat{v}_{z\alpha}}{\partial z_1} + \frac{\partial \hat{v}_{z1}}{\partial z_2} + \nabla \cdot \hat{\mathbf{v}}_{\alpha^2} = 0. \quad (32)$$

The inner and outer expansions of the Navier-Stokes equation (10) governing the radial motion give

$$\frac{\partial \hat{P}_1}{\partial y} = \frac{\partial \hat{P}_\alpha}{\partial y} = 0, \quad (33)$$

$$j\omega \hat{V}_{r\alpha} - \frac{1}{l} \frac{\partial \hat{P}_{\alpha^2}}{\partial y} = 0, \quad (34)$$

$$j\omega \hat{v}_{r1} + \frac{1}{l} \frac{\partial \hat{p}_1}{\partial r} = 0, \quad j\omega \hat{v}_{r\alpha} + \frac{1}{l} \frac{\partial \hat{p}_\alpha}{\partial r} = 0, \quad (35)$$

$$j\omega \hat{v}_{r\alpha^2} - \frac{l^2}{d^2} \frac{\partial^2 \hat{v}_{r1}}{\partial z^2} - \frac{1}{d^2} \left(\frac{4}{3} + \eta_0 \right) \left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \hat{v}_{r1}}{\partial r} \right) - \frac{\hat{v}_{r1}}{r^2} \right] - \frac{1}{d^2 r^2} \frac{\partial^2 \hat{v}_{r1}}{\partial \phi^2} - \frac{l}{d^2} \left(\frac{1}{3} + \eta_0 \right) \frac{\partial^2 \hat{v}_{z1}}{\partial z \partial r} - \frac{1}{d^2 r} \left(\frac{1}{3} + \eta_0 \right) \times \left(\frac{\partial^2 \hat{v}_{\phi 1}}{\partial r \partial \phi} - \frac{1}{r} \frac{\partial \hat{v}_{\phi 1}}{\partial \phi} \right) + \frac{2}{d^2 r^2} \frac{\partial \hat{v}_{\phi 1}}{\partial \phi} + \frac{1}{l} \frac{\partial \hat{p}_1}{\partial r} = 0. \quad (36)$$

Having substituted the expansions (21) into the Navier-Stokes equations for the axial and azimuthal directions (9) and (11) and using Eq. (27), one obtains

$$j\omega \hat{V}_{z1} - \frac{1}{d^2} \frac{\partial^2 \hat{V}_{z1}}{\partial y^2} + \frac{\partial \hat{P}_1}{\partial z} = 0, \quad (37)$$

$$j\omega \hat{V}_{\phi 1} - \frac{1}{d^2} \frac{\partial^2 \hat{V}_{\phi 1}}{\partial y^2} + \frac{1}{l} \frac{\partial \hat{P}_1}{\partial \phi} = 0,$$

$$j\omega \hat{V}_{z\alpha} - \frac{1}{d^2} \frac{\partial^2 \hat{V}_{z\alpha}}{\partial y^2} + \frac{1}{d^2} \frac{\partial \hat{V}_{z1}}{\partial y} + \frac{\partial \hat{P}_\alpha}{\partial z} + \frac{\partial \hat{P}_1}{\partial z_1} = 0,$$

$$j\omega \hat{V}_{\phi\alpha} - \frac{1}{d^2} \frac{\partial^2 \hat{V}_{\phi\alpha}}{\partial y^2} + \frac{1}{d^2} \frac{\partial \hat{V}_{\phi 1}}{\partial y} + \frac{1}{l} \left(\frac{\partial \hat{P}_\alpha}{\partial \phi} + y \frac{\partial \hat{P}_1}{\partial \phi} \right) = 0, \quad (38)$$

$$j\omega \hat{v}_{z1} + \frac{\partial \hat{p}_1}{\partial z} = 0, \quad j\omega \hat{v}_{\phi 1} + \frac{1}{lr} \frac{\partial \hat{p}_1}{\partial \phi} = 0, \quad (39)$$

$$j\omega\hat{v}_{z\alpha} + \frac{\partial\hat{p}_\alpha}{\partial z} + \frac{\partial\hat{p}_1}{\partial z_1} = 0, \quad j\omega\hat{v}_{\phi\alpha} + \frac{1}{lr} \frac{\partial\hat{p}_\alpha}{\partial\phi} = 0, \quad (40)$$

$$j\omega\hat{v}_{z\alpha^2} - \frac{l^2}{d^2} \left(\frac{4}{3} + \eta_0 \right) \frac{\partial^2\hat{v}_{z1}}{\partial z^2} - \frac{1}{d^2 r} \frac{\partial}{\partial r} \left(r \frac{\partial\hat{v}_{z1}}{\partial r} \right) - \frac{1}{d^2 r^2} \frac{\partial^2\hat{v}_{\phi1}}{\partial\phi^2} - \frac{l}{d^2 r} \left(\frac{1}{3} + \eta_0 \right) \left[\frac{\partial}{\partial r} \left(r \frac{\partial\hat{v}_{r1}}{\partial r} \right) + \frac{\partial^2\hat{v}_{\phi1}}{\partial z \partial\phi} \right] + \frac{\partial\hat{p}_{\alpha^2}}{\partial z} + \frac{\partial\hat{p}_\alpha}{\partial z_1} + \frac{\partial\hat{p}_1}{\partial z_2} = 0. \quad (41)$$

$$j\omega\hat{v}_{\phi\alpha^2} - \frac{l^2}{d^2} \frac{\partial^2\hat{v}_{\phi1}}{\partial z^2} - \frac{1}{d^2 r} \left[\frac{\partial}{\partial r} \left(r \frac{\partial\hat{v}_{\phi1}}{\partial r} \right) - \frac{\hat{v}_{\phi1}}{r} \right] - \frac{1}{d^2 r^2} \left(\frac{4}{3} + \eta_0 \right) \frac{\partial^2\hat{v}_{\phi1}}{\partial\phi^2} - \frac{l}{d^2 r} \left(\frac{1}{3} + \eta_0 \right) \frac{\partial^2\hat{v}_{z1}}{\partial z \partial\phi} - \frac{1}{d^2 r^2} \left(\frac{1}{3} + \eta_0 \right) \frac{\partial}{\partial r} \left(r \frac{\partial\hat{v}_{r1}}{\partial r} \right) - \frac{2}{d^2 r^2} \frac{\partial\hat{v}_{r1}}{\partial\phi} + \frac{1}{lr} \frac{\partial\hat{p}_{\alpha^2}}{\partial\phi} = 0.$$

The set of equations following from the inner expansion of Eq. (14) reads

$$\hat{\Psi}_1 = \gamma_0 \hat{P}_1 - \hat{\Theta}_1, \quad \hat{\Psi}_\alpha = \gamma_0 \hat{P}_\alpha - \hat{\Theta}_\alpha, \quad (42)$$

$$\hat{\Psi}_{\alpha^2} = \gamma_0 \hat{P}_{\alpha^2} - \hat{\Theta}_{\alpha^2},$$

which holds for the outer expansion as well. Here, the relationship $G_0 \tilde{\partial}_0 \tilde{\beta}_0 = \gamma_0 - 1$ resulting from Eq. (17) has been used.

The term in the energy equation (13) containing the dissipation function turns out to be of the order $O(M^2\alpha^2)$ in the core region and $O(M^2)$ in the boundary layer. Therefore, the expansions of Eq. (13) take the form

$$j\omega\hat{\Theta}_1 - j\omega(\gamma_0 - 1)\hat{P}_1 - \frac{1}{d^2 \text{Pr}} \frac{\partial^2\hat{\Theta}_1}{\partial y^2} = 0, \quad (43)$$

$$j\omega\hat{\Theta}_\alpha - j\omega(\gamma_0 - 1)\hat{P}_\alpha - \frac{1}{d^2 \text{Pr}} \frac{\partial^2\hat{\Theta}_\alpha}{\partial y^2} + \frac{1}{d^2 \text{Pr}} \frac{\partial\hat{\Theta}_1}{\partial y} = 0, \quad (44)$$

$$\hat{\vartheta}_1 = (\gamma_0 - 1)\hat{p}_1, \quad \hat{\vartheta}_\alpha = (\gamma_0 - 1)\hat{p}_\alpha, \quad (45)$$

$$j\omega\hat{\vartheta}_{\alpha^2} - j\omega(\gamma_0 - 1)\hat{p}_{\alpha^2} - \frac{l^2}{d^2 \text{Pr}} \frac{\partial^2\hat{\vartheta}_1}{\partial z^2} - \frac{1}{d^2 \text{Pr}} \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial\hat{\vartheta}_1}{\partial r} \right) - \frac{1}{d^2 \text{Pr}} \frac{1}{r^2} \frac{\partial^2\hat{\vartheta}_1}{\partial\phi^2} = 0, \quad (46)$$

whereupon substitution of Eqs. (42) into Eqs. (45) yields

$$\hat{p}_1 = \hat{p}_1, \quad \hat{p}_\alpha = \hat{p}_\alpha. \quad (47)$$

Inspection of Eqs. (27), (30), (35), and (39) shows that the leading order pressure perturbations in the core region \hat{p}_1 satisfy the Helmholtz equation

$$\omega^2 \hat{p}_1 + \Delta \hat{p}_1 = 0, \quad (48)$$

together with the boundary condition

$$r=1: \quad \frac{\partial\hat{p}_1}{\partial r} = 0. \quad (49)$$

The solution is given by the modal decomposition

$$\hat{p}_1 = \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} \left(\hat{C}_{1nm1}(z_1, z_2) \frac{J_m(\gamma_{nm}r)}{J_m(\gamma_{nm})} e^{-j(\hat{k}_{nm}z + m\phi)} + \hat{C}_{2nm1}(z_1, z_2) \frac{J_m(\gamma_{nm}r)}{J_m(\gamma_{nm})} e^{j(\hat{k}_{nm}z + m\phi)} \right), \quad (50)$$

where the parameters $\gamma_{nm} = \gamma_{n(-m)}$, $n=0,1,2,\dots$, are the zeros of the first derivatives of the Bessel functions $\partial J_m(\xi)/\partial\xi$, with $\gamma_{00}=0$, and the quantities

$$\hat{k}_{nm} = \hat{k}_{n(-m)} = \begin{cases} \sqrt{\omega^2 - \gamma_{nm}^2/l^2}, & \omega > \gamma_{nm}/l, \\ -j\sqrt{\gamma_{nm}^2/l^2 - \omega^2}, & \omega < \gamma_{nm}/l, \end{cases} \quad (51)$$

are, respectively, the axial wave numbers for the propagating and evanescent modes.

Combining Eqs. (31), (35), (39), and (40) yields the inhomogeneous Helmholtz equation for the second-order term \hat{p}_α :

$$\omega^2 \hat{p}_\alpha + \Delta \hat{p}_\alpha = -2 \frac{\partial^2 \hat{p}_1}{\partial z \partial z_1}. \quad (52)$$

Upon applying Van Dyke's matching rules to the expressions for \hat{V}_{z1} , \hat{v}_{z1} , $\hat{V}_{\phi1}$, $\hat{v}_{\phi1}$, $\hat{\Theta}_1$, and $\hat{\vartheta}_1$ resulting from Eqs. (37), (39), (43), (45), and (19), one obtains

$$\hat{V}_{z1} = \hat{v}_{z1}|_{r=1} (1 - e^{-(1+j)dy\sqrt{\omega/2}}),$$

$$\hat{V}_{\phi1} = \hat{v}_{\phi1}|_{r=1} (1 - e^{-(1+j)dy\sqrt{\omega/2}}), \quad (53)$$

$$\hat{\Theta}_1 = \hat{p}_1|_{r=1} (\gamma_0 - 1) (1 - e^{-(1+j)dy\sqrt{\omega \text{Pr}/2}}).$$

It then follows from Eqs. (28) and (42) that

$$\frac{\hat{V}_{r\alpha}}{l} = \hat{p}_1|_{r=1} \left[j\omega y + \frac{1+j}{d} \sqrt{\frac{\omega}{2}} \frac{\gamma_0 - 1}{\sqrt{\text{Pr}}} (1 - e^{-(1+j)dy\sqrt{\omega \text{Pr}/2}}) \right] + \left(\frac{\partial^2 \hat{p}_1}{\partial z^2} + \frac{1}{l^2} \frac{\partial^2 \hat{p}_1}{\partial\phi^2} \right) \Big|_{r=1} \times \frac{1}{\omega^2} \left[j\omega y - \frac{1+j}{d} \sqrt{\frac{\omega}{2}} (1 - e^{-(1+j)dy\sqrt{\omega/2}}) \right]. \quad (54)$$

As a consequence, the matching principle gives the boundary condition for the outer expansion term \hat{p}_α in the form

$$r=1: \quad \frac{\partial\hat{p}_\alpha}{\partial r} = l^2 \frac{1-j}{d} \omega \sqrt{\frac{\omega}{2}} \left[\hat{p}_1 \frac{\gamma_0 - 1}{\sqrt{\text{Pr}}} - \left(\frac{\partial^2 \hat{p}_1}{\partial z^2} + \frac{1}{l^2} \frac{\partial^2 \hat{p}_1}{\partial\phi^2} \right) \frac{1}{\omega^2} \right]. \quad (55)$$

The solution \hat{p}'_α to the homogeneous part of Eq. (52) subject to the homogeneous boundary conditions at $r=1$ assumes the same form as the expression for \hat{p}_1 from Eq. (50). However, the functions \hat{C}_{1nm1} and \hat{C}_{2nm1} have to be replaced with $\hat{C}_{1nm\alpha}$ and $\hat{C}_{2nm\alpha}$, respectively. In order to determine a particular solution \hat{p}''_α such that $\hat{p}_\alpha = \hat{p}'_\alpha + \hat{p}''_\alpha$ the ansatz

$$\begin{aligned} \hat{p}''_{\alpha} = & \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} \left[\left(\hat{f}_{1nm\alpha}(z, z_1, z_2) \frac{J_m(\gamma_{nm}r)}{J_m(\gamma_{nm})} \right. \right. \\ & + \hat{g}_{nm\alpha}(r) \hat{C}_{1nm1}(z_1, z_2) \left. \right) e^{-j(\hat{k}_{nm}z + m\phi)} \\ & + \left(\hat{f}_{2nm\alpha}(z, z_1, z_2) \frac{J_m(\gamma_{nm}r)}{J_m(\gamma_{nm})} \right. \\ & \left. \left. + \hat{g}_{nm\alpha}(r) \hat{C}_{2nm1}(z_1, z_2) \right) e^{j(\hat{k}_{nm}z + m\phi)} \right] \end{aligned} \quad (56)$$

is made, which is required to be bounded at $r=0$ and to fulfill the condition (55). Furthermore, the functional dependence on the coordinate r of the terms that are generated by the left-hand side of Eq. (52) for each mode (n, m) upon substituting Eq. (56) must be given by the radial eigenfunctions $J_m(\gamma_{nm}r)$. Hence, these requirements imply that

$$\begin{aligned} \hat{g}_{00\alpha} = & \frac{l^2 \omega^2 \hat{F}_{00}}{2} r^2, \quad \hat{g}_{nm\alpha} = l^2 \hat{k}_{nm}^2 \hat{F}_{nm} \frac{r J_{m+1}(\gamma_{nm}r)}{\gamma_{nm} J_m(\gamma_{nm})}, \\ & (n, m) \neq (0, 0), \end{aligned} \quad (57)$$

with the parameters \hat{F}_{nm} defined as

$$\hat{F}_{nm} = \frac{1-j}{d \sqrt{2\omega(1-m^2/\gamma_{nm}^2)}} \left(1 + \frac{m^2}{l^2 \hat{k}_{nm}^2} + \frac{\omega^2}{\hat{k}_{nm}^2} \frac{\gamma_0 - 1}{\sqrt{\text{Pr}}} \right). \quad (58)$$

Substitution of the expression (56) into Eq. (52) then results in

$$\begin{aligned} & \frac{\partial^2 \hat{f}_{1nm\alpha}}{\partial z^2} - 2j\hat{k}_{nm} \frac{\partial \hat{f}_{1nm\alpha}}{\partial z} \\ & = -2\hat{F}_{nm} \hat{k}_{nm}^2 \hat{C}_{1nm1} + 2j\hat{k}_{nm} \frac{\partial \hat{C}_{1nm1}}{\partial z_1}, \\ & \frac{\partial^2 \hat{f}_{2nm\alpha}}{\partial z^2} + 2j\hat{k}_{nm} \frac{\partial \hat{f}_{2nm\alpha}}{\partial z} \\ & = -2\hat{F}_{nm} \hat{k}_{nm}^2 \hat{C}_{2nm1} - 2j\hat{k}_{nm} \frac{\partial \hat{C}_{2nm1}}{\partial z_1}. \end{aligned} \quad (59)$$

The forcing terms on the right-hand sides are resonant and would produce secular terms in the functions $\hat{f}_{1nm\alpha}$ and $\hat{f}_{2nm\alpha}$. Therefore, they must be annihilated, yielding

$$\begin{aligned} \hat{C}_{1nm1}(z_1, z_2) & \rightarrow \hat{C}_{1nm1}(z_2) e^{-j\hat{k}_{nm} \hat{F}_{nm} z_1}, \\ \hat{C}_{2nm1}(z_1, z_2) & \rightarrow \hat{C}_{2nm1}(z_2) e^{j\hat{k}_{nm} \hat{F}_{nm} z_1}, \end{aligned} \quad (60)$$

whereupon the functions $\hat{f}_{1nm\alpha}$ and $\hat{f}_{2nm\alpha}$ can simply be set to zero since any other solution would lead to expressions for \hat{p}''_{α} that could be incorporated into the homogeneous solution \hat{p}'_{α} . Please note that the solvability conditions (60) could also have been derived by employing the so-called *Fredholm alternative* (see, for example, Ref. 29). This method will be used in Sec. IV.

The quantities $\mathfrak{J}(-\omega \hat{F}_{00})$ and $\mathfrak{J}(-\hat{k}_{nm} \hat{F}_{nm})$, $(n, m) \neq (0, 0)$, turn out to be, respectively, the leading order decay rate of the fundamental mode and the leading order decay rates of the higher order modes due to boundary layer attenuation. This is in accordance with Beatty's results⁷ obtained by using the concept of boundary layer admittance.

In order to calculate the corresponding expressions for the third-order problem, a procedure very similar to that used before is performed: A combination of Eqs. (32), (36), (39)–(41), (45), and (46) and multiple application of Eq. (48) shows that

$$\hat{v}_{z\alpha^2} = -\frac{l^2}{d^2} \frac{\partial \hat{p}_1}{\partial z} \left(\frac{4}{3} + \eta_0 \right) + \frac{j}{\omega} \left(\frac{\partial \hat{p}_1}{\partial z_2} + \frac{\partial \hat{p}_{\alpha}}{\partial z_1} + \frac{\partial \hat{p}_{\alpha^2}}{\partial z} \right) \quad (61)$$

and that \hat{p}_{α^2} has to satisfy the inhomogeneous equation

$$\begin{aligned} \omega^2 \hat{p}_{\alpha^2} + \Delta \hat{p}_{\alpha^2} = & -2 \frac{\partial^2 \hat{p}_1}{\partial z \partial z_2} - \frac{\partial^2 \hat{p}_1}{\partial z_1^2} + l^2 \frac{j\omega^3}{d^2} \\ & \times \hat{p}_1 \left(\frac{4}{3} + \eta_0 + \frac{\gamma_0 - 1}{\text{Pr}} \right) - 2 \frac{\partial^2 \hat{p}_{\alpha}}{\partial z \partial z_1}, \end{aligned} \quad (62)$$

where the third term on the right-hand side obviously incorporates the effects of heat conduction and viscosity in the core region associated with the propagation of the leading order pressure fluctuations \hat{p}_1 . Moreover, Eqs. (38) and (44) are solved for the axial and the azimuthal velocity and the temperature in the boundary layer, leading to

$$\begin{aligned} \hat{V}_{z\alpha} = & \hat{v}_{z\alpha}|_{r=1} (1 - e^{-(1+j)y\sqrt{\omega/2}}) \\ & - \hat{v}_{z1}|_{r=1} \frac{y}{2} e^{-(1+j)y\sqrt{\omega/2}}, \\ \hat{V}_{\phi\alpha} = & \hat{v}_{\phi\alpha}|_{r=1} (1 - e^{-(1+j)y\sqrt{\omega/2}}) \\ & + \hat{v}_{\phi1}|_{r=1} y (1 - \frac{1}{2} e^{-(1+j)y\sqrt{\omega/2}}), \end{aligned} \quad (63)$$

$$\begin{aligned} \hat{\Theta}_{\alpha} = & \hat{p}_{\alpha}|_{r=1} (\gamma_0 - 1) (1 - e^{-(1+j)y\sqrt{\omega \text{Pr}/2}}) \\ & - \hat{p}_1|_{r=1} (\gamma_0 - 1) \frac{y}{2} e^{-(1+j)y\sqrt{\omega \text{Pr}/2}}. \end{aligned}$$

After substituting for $\hat{\Psi}_{\alpha}$, \hat{V}_{z1} , $\hat{V}_{\phi1}$, $\hat{V}_{r\alpha}$, $\hat{V}_{z\alpha}$, and $\hat{V}_{\phi\alpha}$ from Eqs. (42), (53), (54), and (63), respectively, integration of Eq. (29) with respect to the inner coordinate y gives

$$\begin{aligned}
\frac{\hat{V}_{r\alpha^2}}{l} = & \hat{p}_1|_{r=1} \frac{1}{2} \left[j\omega y^2 - \frac{\gamma_0 - 1}{d^2 \text{Pr}} (1 - e^{-(1+j)dy\sqrt{\omega \text{Pr}/2}}) + \frac{1+j}{d} \sqrt{\frac{\omega}{2}} \frac{\gamma_0 - 1}{\sqrt{\text{Pr}}} y (2 - e^{-(1+j)dy\sqrt{\omega \text{Pr}/2}}) \right] \\
& + \frac{\partial^2 \hat{p}_1}{\partial z^2} \Big|_{r=1} \frac{1}{2\omega^2} \left[j\omega y^2 + \frac{1}{d^2} (1 - e^{-(1+j)dy\sqrt{\omega/2}}) - \frac{1+j}{d} \sqrt{\frac{\omega}{2}} y (2 - e^{-(1+j)dy\sqrt{\omega/2}}) \right] \\
& + \frac{1}{l^2} \frac{\partial^2 \hat{p}_1}{\partial \phi^2} \Big|_{r=1} \frac{1}{2\omega^2} \left[3j\omega y^2 - \frac{1}{d^2} (1 - e^{-(1+j)dy\sqrt{\omega/2}}) - \frac{1+j}{d} \sqrt{\frac{\omega}{2}} y \left(2 - \frac{3}{2} e^{-(1+j)dy\sqrt{\omega/2}} \right) \right] \\
& + \hat{p}_\alpha|_{r=1} \left[j\omega y + \frac{1+j}{d} \sqrt{\frac{\omega}{2}} \frac{\gamma_0 - 1}{\sqrt{\text{Pr}}} (1 - e^{-(1+j)dy\sqrt{\omega \text{Pr}/2}}) \right] + \left(\frac{\partial^2 \hat{p}_\alpha}{\partial z^2} + \frac{1}{l^2} \frac{\partial^2 \hat{p}_\alpha}{\partial \phi^2} + 2 \frac{\partial^2 \hat{p}_1}{\partial z \partial z_1} \right) \Big|_{r=1} \\
& \times \frac{1}{\omega^2} \left[j\omega y - \frac{1+j}{d} \sqrt{\frac{\omega}{2}} (1 - e^{-(1+j)dy\sqrt{\omega/2}}) \right]. \tag{64}
\end{aligned}$$

With the matching rules applied to the inner and outer expansions of the radial velocity, the boundary condition

$$\begin{aligned}
r=1: \quad \frac{\partial \hat{p}_{\alpha^2}}{\partial r} = & l^2 \frac{j\omega}{2d^2} \left[\hat{p}_1 \frac{\gamma_0 - 1}{\text{Pr}} - \left(\frac{\partial^2 \hat{p}_1}{\partial z^2} - \frac{1}{l^2} \frac{\partial^2 \hat{p}_1}{\partial \phi^2} \right) \frac{1}{\omega^2} \right] \\
& + l^2 \frac{1-j}{d} \omega \sqrt{\frac{\omega}{2}} \left[\hat{p}_\alpha \frac{\gamma_0 - 1}{\sqrt{\text{Pr}}} - \left(\frac{\partial^2 \hat{p}_\alpha}{\partial z^2} \right. \right. \\
& \left. \left. + \frac{1}{l^2} \frac{\partial^2 \hat{p}_\alpha}{\partial \phi^2} + 2 \frac{\partial^2 \hat{p}_1}{\partial z \partial z_1} \right) \frac{1}{\omega^2} \right] \tag{65}
\end{aligned}$$

is derived.

A particular solution \hat{p}''_{α^2} to Eq. (62) can be found by employing the ansatz

$$\begin{aligned}
\hat{p}''_{\alpha^2} = & \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} \left[\left(\hat{f}_{1nm\alpha^2}(z, z_1, z_2) \frac{J_m(\gamma_{nm}r)}{J_m(\gamma_{nm})} \right. \right. \\
& + \hat{g}_{nm\alpha^2}(r) \hat{C}_{1nm1}(z_2) e^{-j\hat{F}_{nm}\hat{k}_{nm}z_1} \\
& + \hat{h}_{nm\alpha^2}(r) \hat{C}_{1nm\alpha}(z_1, z_2) \left. \right) e^{-j(\hat{k}_{nm}z + m\phi)} \\
& + \left(\hat{f}_{2nm\alpha^2}(z, z_1, z_2) \frac{J_m(\gamma_{nm}r)}{J_m(\gamma_{nm})} \right. \\
& + \hat{g}_{nm\alpha^2}(r) \hat{C}_{2nm1}(z_2) e^{j\hat{F}_{nm}\hat{k}_{nm}z_1} \\
& \left. \left. + \hat{h}_{nm\alpha^2}(r) \hat{C}_{2nm\alpha}(z_1, z_2) \right) e^{j(\hat{k}_{nm}z - m\phi)} \right] \tag{66}
\end{aligned}$$

such that the boundary condition (65) is satisfied and the operators on the left-hand side of Eq. (62) produce terms that have the same functional dependence on r as the terms on the right-hand side. Consequently, one obtains

$$\hat{g}_{00\alpha^2} = \frac{l^2 \omega^2}{2} \left(\hat{G}_{00} + \frac{\omega^2 \hat{F}_{00}^2}{4} \right) r^2 + \frac{l^4 \omega^4 \hat{F}_{00}^2}{16} r^4,$$

$$\begin{aligned}
\hat{g}_{nm\alpha^2} = & l^2 \hat{k}_{nm}^2 \left[\hat{G}_{nm} + \left(m + \frac{1}{1 - m^2/\gamma_{nm}^2} \right) \right. \\
& \times \frac{l^2 \hat{k}_{nm}^2 \hat{F}_{nm}^2}{\gamma_{nm}^2} \left. \frac{r J_{m+1}(\gamma_{nm}r)}{\gamma_{nm} J_m(\gamma_{nm})} \right. \\
& \left. - \frac{l^4 \hat{k}_{nm}^4 \hat{F}_{nm}^2}{2} \frac{r^2 J_m(\gamma_{nm}r)}{\gamma_{nm}^2 J_m(\gamma_{nm})} \right], \quad (n, m) \neq (0, 0), \tag{67}
\end{aligned}$$

$$\begin{aligned}
\hat{h}_{00\alpha^2} = & \frac{l^2 \omega^2 \hat{F}_{00}}{2} r^2, \quad \hat{h}_{nm\alpha^2} = l^2 \hat{k}_{nm}^2 \hat{F}_{nm} \frac{r J_{m+1}(\gamma_{nm}r)}{\gamma_{nm} J_m(\gamma_{nm})}, \\
& (n, m) \neq (0, 0),
\end{aligned}$$

where

$$\begin{aligned}
\hat{G}_{nm} = & - \frac{j}{2d^2 \omega (1 - m^2/\gamma_{nm}^2)^2} \left[3 + \frac{m^2}{\gamma_{nm}^2} + \frac{m^2}{l^2 \hat{k}_{nm}^2} \left(5 - \frac{m^2}{\gamma_{nm}^2} \right) \right. \\
& \left. - \frac{\omega^2}{\hat{k}_{nm}^2} (\gamma_0 - 1) \left(\left(1 - \frac{m^2}{\gamma_{nm}^2} \right) \frac{1}{\text{Pr}} - \frac{4}{\sqrt{\text{Pr}}} \right) \right]. \tag{68}
\end{aligned}$$

Substitution of Eq. (56) into Eq. (62) then yields

$$\begin{aligned}
& \frac{\partial^2 \hat{f}_{1nm\alpha^2}}{\partial z^2} - 2j\hat{k}_{nm} \frac{\partial \hat{f}_{1nm\alpha^2}}{\partial z} \\
& = \left(-2\hat{H}_{nm} \hat{k}_{nm}^2 \hat{C}_{1nm1} + 2j\hat{k}_{nm} \frac{\partial \hat{C}_{1nm1}}{\partial z_2} \right) e^{-j\hat{F}_{nm}\hat{k}_{nm}z_1} \\
& \quad - 2\hat{F}_{nm} \hat{k}_{nm}^2 \hat{C}_{1nm\alpha} + 2j\hat{k}_{nm} \frac{\partial \hat{C}_{1nm\alpha}}{\partial z_1},
\end{aligned}$$

$$\begin{aligned} & \frac{\partial^2 \hat{f}_{2nm\alpha^2}}{\partial z^2} + 2j\hat{k}_{nm} \frac{\partial \hat{f}_{2nm\alpha^2}}{\partial z} \\ &= \left(-2\hat{H}_{nm}\hat{k}_{nm}^2 \hat{C}_{1nm1} - 2j\hat{k}_{nm} \frac{\partial \hat{C}_{2nm1}}{\partial z_2} \right) e^{j\hat{F}_{nm}\hat{k}_{nm}z_1} \\ & \quad - 2\hat{F}_{nm}\hat{k}_{nm}^2 \hat{C}_{2nm\alpha} - 2j\hat{k}_{nm} \frac{\partial \hat{C}_{2nm\alpha}}{\partial z_1}, \end{aligned} \quad (69)$$

where

$$\begin{aligned} \hat{H}_{00} &= \hat{G}_{00} - \frac{\hat{F}_{00}^2}{2} + \frac{l^2 \omega^2 \hat{F}_{00}^2}{4} - l^2 \frac{j\omega}{2d^2} \left(\frac{4}{3} + \eta_0 + \frac{\gamma_0 - 1}{\text{Pr}} \right) \\ &= -\frac{j}{d^2 \omega} \left[1 + \frac{\gamma_0 - 1}{\sqrt{\text{Pr}}} \left(1 - \frac{\gamma_0}{2\sqrt{\text{Pr}}} \right) \right] \\ & \quad - l^2 \frac{j\omega}{2d^2} \left[\frac{11}{6} + \eta_0 + \frac{\gamma_0 - 1}{2\sqrt{\text{Pr}}} \left(2 + \frac{1 + \gamma_0}{\sqrt{\text{Pr}}} \right) \right], \end{aligned} \quad (70)$$

$$\begin{aligned} \hat{H}_{nm} &= \hat{G}_{nm} - \frac{\hat{F}_{nm}^2}{2} + \frac{l^2 \hat{k}_{nm}^2 \hat{F}_{nm}^2}{\gamma_{nm}^2} - \frac{m^2}{\gamma_{nm}^2 - m^2} - l^2 \frac{j\omega^2}{2d^2 \hat{k}_{nm}^2} \\ & \quad \times \left(\frac{4}{3} + \eta_0 + \frac{\gamma_0 - 1}{\text{Pr}} \right), \quad (n, m) \neq (0, 0). \end{aligned}$$

As before, the right-hand sides of Eqs. (69) have to be annihilated in order to rule out secular terms in the functions $\hat{f}_{1nm\alpha^2}$ and $\hat{f}_{2nm\alpha^2}$. The resulting solvability conditions thus read

$$\begin{aligned} & 2\hat{F}_{nm}\hat{k}_{nm}^2 \hat{C}_{1nm\alpha} - 2j\hat{k}_{nm} \frac{\partial \hat{C}_{1nm\alpha}}{\partial z_1} \\ &= \left(-2\hat{H}_{nm}\hat{k}_{nm}^2 \hat{C}_{1nm1} + 2j\hat{k}_{nm} \frac{\partial \hat{C}_{1nm1}}{\partial z_2} \right) e^{-j\hat{F}_{nm}\hat{k}_{nm}z_1}, \end{aligned} \quad (71)$$

$$\begin{aligned} & 2\hat{F}_{nm}\hat{k}_{nm}^2 \hat{C}_{2nm\alpha} + 2j\hat{k}_{nm} \frac{\partial \hat{C}_{2nm\alpha}}{\partial z_1} \\ &= \left(-2\hat{H}_{nm}\hat{k}_{nm}^2 \hat{C}_{2nm1} - 2j\hat{k}_{nm} \frac{\partial \hat{C}_{2nm1}}{\partial z_2} \right) e^{j\hat{F}_{nm}\hat{k}_{nm}z_1}, \end{aligned}$$

and, furthermore, $\hat{f}_{1nm\alpha^2}$ and $\hat{f}_{2nm\alpha^2}$ can be set to zero. The right-hand sides of Eqs. (71) are again identified as resonant forcing terms. Their elimination leads to

$$\hat{C}_{1nm1}(z_2) e^{-j\hat{F}_{nm}\hat{k}_{nm}z_1} \rightarrow \hat{C}_{1nm1} e^{-j\hat{k}_{nm}(\hat{F}_{nm}z_1 + \hat{H}_{nm}z_2)}, \quad (72)$$

$$\hat{C}_{2nm1}(z_2) e^{j\hat{F}_{nm}\hat{k}_{nm}z_1} \rightarrow \hat{C}_{2nm1} e^{j\hat{k}_{nm}(\hat{F}_{nm}z_1 + \hat{H}_{nm}z_2)},$$

and, finally,

$$\hat{C}_{1nm\alpha}(z_1, z_2) \rightarrow \hat{C}_{1nm\alpha}(z_2) e^{-j\hat{k}_{nm}\hat{F}_{nm}z_1}, \quad (73)$$

$$\hat{C}_{2nm\alpha}(z_1, z_2) \rightarrow \hat{C}_{2nm\alpha}(z_2) e^{j\hat{k}_{nm}\hat{F}_{nm}z_1}.$$

A. Results

Combining the expressions derived so far yields the following solution for the sound pressure, which is valid in the complete spatial range defined by $z = O(\alpha^{-1})$:

$$\begin{aligned} \hat{p}_1 &= \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} \left(\hat{C}_{1nm1} \frac{J_m(\gamma_{nm}r)}{J_m(\gamma_{nm})} \right. \\ & \quad \times e^{-j[\hat{k}_{nm}(z + \hat{F}_{nm}z_1 + \hat{H}_{nm}z_2) + m\phi]} + \hat{C}_{2nm1} \frac{J_m(\gamma_{nm}r)}{J_m(\gamma_{nm})} \\ & \quad \times e^{j[\hat{k}_{nm}(z + \hat{F}_{nm}z_1 + \hat{H}_{nm}z_2) + m\phi]} \Big), \end{aligned} \quad (74)$$

$$\begin{aligned} \hat{p}_\alpha &= \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} \left[\left(\hat{C}_{1nm\alpha} \frac{J_m(\gamma_{nm}r)}{J_m(\gamma_{nm})} \right. \right. \\ & \quad \left. \left. + \hat{C}_{1nm1} \hat{g}_{nm\alpha}(r) \right) e^{-j[\hat{k}_{nm}(z + \hat{F}_{nm}z_1) + m\phi]} \right. \\ & \quad \left. + \left(\hat{C}_{2nm\alpha} \frac{J_m(\gamma_{nm}r)}{J_m(\gamma_{nm})} + \hat{C}_{2nm1} \hat{g}_{nm\alpha}(r) \right) \right. \\ & \quad \left. \times e^{j[\hat{k}_{nm}(z + \hat{F}_{nm}z_1) + m\phi]} \right], \end{aligned} \quad (75)$$

where \hat{F}_{nm} and \hat{H}_{nm} are the quantities already introduced in Eqs. (58) and (70), respectively; the function $\hat{g}_{nm\alpha}$ is given by Eq. (57). Please note that the next higher order correction,

$$\begin{aligned} \hat{p}_{\alpha^2} &= \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} \left(\hat{C}_{1nm\alpha^2}(z_1, z_2) \frac{J_m(\gamma_{nm}r)}{J_m(\gamma_{nm})} e^{-j(\hat{k}_{nm}z + m\phi)} \right. \\ & \quad \left. + \hat{C}_{2nm\alpha^2}(z_1, z_2) \frac{J_m(\gamma_{nm}r)}{J_m(\gamma_{nm})} e^{j(\hat{k}_{nm}z + m\phi)} \right) + \hat{p}_{\alpha^2}''', \end{aligned} \quad (76)$$

with \hat{p}_{α^2}''' given by Eq. (66), is valid over ranges of the order $z = O(1)$ only because the dependence of the functions $\hat{C}_{1nm\alpha^2}$ and $\hat{C}_{2nm\alpha^2}$ on the scales z_1 and z_2 remains undetermined at the level of approximation to be reached here.

In order to define the propagation parameters $\hat{\Gamma}_{nm}$, each mode has to be treated separately:

$$\hat{\Gamma}_{nm} = \sqrt{\frac{1}{\hat{p}_{nms} \hat{v}_{nmz}} \frac{\partial \hat{p}_{nms}}{\partial z} \frac{\partial \hat{v}_{nmz}}{\partial z}}, \quad (77)$$

where \hat{p}_{nms} and \hat{v}_{nmz} denote the sound pressures and the axial velocities associated with the modes $n=0,1,2,\dots$, $m = \dots, -1, 0, 1, \dots$. Inspection of Eqs. (39), (40), and (61) shows that \hat{v}_{nmz} can be expanded as

$$\hat{v}_{nmz} = \frac{j}{\omega} \frac{\partial \hat{p}_{nms}}{\partial z} - \alpha^2 \frac{l^2}{d^2} \frac{\partial \hat{p}_{nms}}{\partial z} \left(\frac{4}{3} + \eta_0 \right) + \dots, \quad (78)$$

keeping in mind that the partial derivatives with respect to z have to be understood in the sense of derivatives with respect to the three length scales introduced by Eq. (24). Since, after substitution of Eq. (78) into Eq. (77) and expansion with respect to α , all terms proportional to $(\frac{4}{3} + \eta_0)$ cancel out, the definition of the propagation parameters can be equivalently written as

$$\hat{\Gamma}_{nm} = \sqrt{\frac{1}{\hat{p}_{nms}} \frac{\partial^2 \hat{p}_{nms}}{\partial z^2} + \dots} \quad (79)$$

Once the derivative $\partial^2/\partial z^2$ has been replaced with derivatives with respect to the multiple scales $z, z_1,$ and z_2 , application of Eqs. (74)–(76) implicates that the parameters $\hat{\Gamma}_{nm}$ are given by the relationship

$$\hat{\Gamma}_{nm} = \sqrt{-\hat{k}_{nm}^2 [1 + 2\alpha \hat{F}_{nm} + \alpha^2 (2\hat{H}_{nm} + \hat{F}_{nm}^2)] + \dots} \quad (80)$$

Interestingly, \hat{H}_{00} and thus $\hat{\Gamma}_{00}$ cannot be derived from \hat{H}_{nm} and $\hat{\Gamma}_{nm}$ for the higher order modes simply by setting $\hat{k}_{nm} = \omega, \gamma_{nm} = 0,$ and $m = 0.$

If $|\hat{k}_{nm}| = O(1),$ the result (80) simplifies to

$$\hat{\Gamma}_{nm} = j\hat{k}_{nm} + \alpha j\hat{k}_{nm}\hat{F}_{nm} + \alpha^2 j\hat{k}_{nm}\hat{H}_{nm} + \dots \quad (81)$$

In contrast to that, inspection of the definitions (58) and (70) shows that in the limiting case of a mode $(n, m) \neq (0, 0)$ being driven so close to its cutoff frequency that $|\hat{k}_{nm}| \ll 1,$ the orders of magnitude of the different terms appearing in Eq. (80) may change completely. Moreover, the solvability conditions resulting from Eqs. (59) and (69) turn out to be singular at the cutoff frequencies $\omega = \gamma_{nm}/l,$ predicting that in the limit as $\omega \rightarrow \gamma_{nm}/l, \hat{k}_{nm} \rightarrow 0,$ the attenuation of the mode $(n, m) = (\bar{n}, \bar{m})$ takes place over a much shorter spatial range than that defined by $z_1 = O(1).$ As explicated in Scheichl,⁶ however, the expression for $\hat{\Gamma}_{nm}$ from Eq. (80) retains its validity even for those driving frequencies that are very close to the cutoff frequency of this mode. Expanding ω as $\omega = \omega_{nm} + \alpha\omega_\alpha,$ with $\omega_{nm} = \gamma_{nm}/l,$ one obtains the following expression for the propagation parameter

$$\hat{\Gamma}_{nm} = \alpha^{1/2} \hat{\Gamma}_{nm\alpha^{1/2}} + \alpha^{3/2} \hat{\Gamma}_{nm\alpha^{3/2}} + \dots, \quad (82)$$

where

$$\begin{aligned} \hat{\Gamma}_{nm\alpha^{1/2}} &= j(2\hat{F}_{nm} + 2\omega_{nm}\omega_\alpha)^{1/2}, \\ \hat{\Gamma}_{nm\alpha^{3/2}} &= \frac{j}{2}(2\hat{F}_{nm} + 2\omega_{nm}\omega_\alpha)^{1/2} \left(\frac{\omega_\alpha}{2\omega_{nm}} \right. \\ &\quad \left. + \frac{\omega_\alpha \hat{F}_{nm} + \sqrt{2\omega_{nm}\omega_\alpha}[(1-j)/d] + \hat{K}_{nm}}{\hat{F}_{nm} + \omega_{nm}\omega_\alpha} \right), \quad (83) \end{aligned}$$

$$\hat{F}_{nm} = \lim_{\omega \rightarrow \omega_{nm}} k_{nm}^2 \hat{F}_{nm}, \quad \hat{K}_{nm} = \lim_{\omega \rightarrow \omega_{nm}} k_{nm}^2 \left(\hat{H}_{nm} + \frac{\hat{F}_{nm}^2}{2} \right),$$

which is valid for all frequencies ω arbitrarily near the cutoff frequency $\gamma_{nm}/l,$ i.e., even for $\omega_\alpha = 0.$ When truncated after the first expansion term, this solution agrees perfectly with the approximation derived by Hudde⁹ using the concept of boundary layer admittance.

Another interesting property of the general expression for the propagation parameters from Eq. (80) is that its formal representation would have remained unchanged if a perfect gas had been considered since the only quantities de-

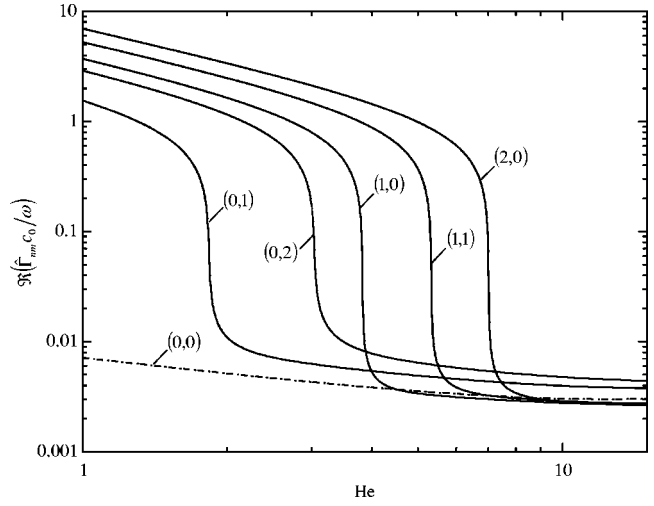


FIG. 2. Graphs of $\Re(\hat{\Gamma}_{nm}c_0/\omega)$ as functions of He in double logarithmic scale, $R = 0.001$ m, i.e., $St = \sqrt{\text{He Re}_R}$ with $\text{Re}_R = c_0 R \rho_0 / \mu_0;$ $\text{Pr} = 0.707, \gamma_0 = 1.402, \eta_0 / \mu_0 = 0.6, c_0 \rho_0 / \mu_0 = 221.3732 \times 10^5 \text{ m}^{-1}$ (air at 300 K and 10^5 Pa); $\gamma_{00} = 0, \gamma_{01} = 1.84118, \gamma_{02} = 3.05424, \gamma_{10} = 3.83171, \gamma_{11} = 5.33144,$ and $\gamma_{20} = 7.01559.$

pending on the thermodynamic equilibrium state of the fluid that enter the definitions (58) and (70) of \hat{F}_{nm} and \hat{H}_{nm} are the parameters $\gamma_0 = \tilde{C}_{p0} / \tilde{C}_{v0}, \text{Pr},$ and $\eta_0.$

The dimensional form of Eq. (80) is given in the Appendix and can be compared with results given in the literature: Kergomard *et al.*³⁰ calculated the propagation parameters for the higher order modes, starting from a generalized dispersion equation (see also Refs. 8 and 31). However, in the intermediate steps that then followed, only boundary layer effects were taken into account [see Eq. (2) in Ref. 30] and some other simplifying hypotheses were applied. Thus, their solutions, i.e., Eqs. (7) and (8) in Ref. 30, differ quite significantly from the expressions for the propagation parameters $\hat{\Gamma}_{nm}$ presented in Eqs. (80) and (A1). When, though, the quantity X in that paper is recalculated without using the simplifications concerning the involved quantities $D_a, D_{v0},$ and D_{h0} and expanded with respect to $St^{-1},$ one obtains a solution for the propagation parameter which is in complete accordance with the results from Eq. (80), provided that the last bracket term appearing in the definition (70) of \hat{H}_{nm} is formally omitted. However, it should be kept in mind that in the case of $l = O(1),$ the terms in \hat{H}_{nm} arising from boundary layer attenuation and the viscous and thermal effects in the core region are of the same order of magnitude. Therefore, they cannot be subject to *a posteriori* changes within the framework of an asymptotically correct analysis.

In order to illustrate the behavior of the modal damping rates when the driving frequency is successively increased, it seems reasonable to scale the dimensional propagation parameters from Eqs. (A1) and (A2) with ω/c_0 since the thus-obtained quantities can be expressed in terms of the parameters $St, \text{He}, \text{Pr}, \gamma_0,$ and $\eta_0/\mu_0.$ The influence of the viscothermal damping mechanisms in the core region on $\hat{\Gamma}_{00}$ and the propagation parameters of the radial modes has already been discussed in Ref. 6. In extension to that, in Fig. 2 the graphs of $\Re(\hat{\Gamma}_{nm}c_0/\omega)$ for all symmetrical and asym-

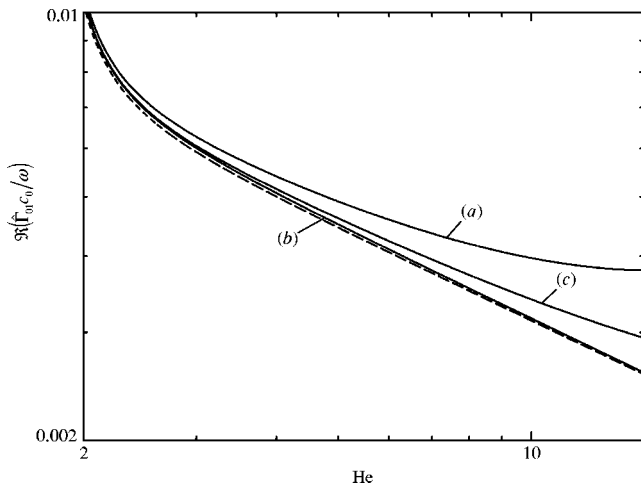


FIG. 3. Comparison of the results for $\Re(\hat{\Gamma}_{01}c_0/\omega)$ obtained from different theories solid lines: (a) complete asymptotic result from Eq. (A1), (b) wide tube solution Eq. (8) from Kergomard *et al.*,³⁰ (c) corrected wide tube solution based on the dispersion relation (2) in Ref. 30, this corresponds to Eq. (A1) without the core region terms. Dashed line: Beatty's solution,⁷ i.e., Eq. (A1) truncated after the first correction term.

metrical modes (n, m) with $n + |m| \leq 2$ are depicted as functions of the Helmholtz number $He = O(1)$. Again, the (large) Stokes number has been eliminated using the relationship $St = \sqrt{He} Re_R$, where $Re_R = c_0 R \rho_0 / \mu_0 \sim Re \gg 1$ denotes the radial Reynolds number. The results plotted in this figure ($R = 0.001$ m) clearly show that the sudden drops of the damping rates when the frequency is increased such that a new mode becomes propagational are in conformity with Eq. (82): $\Re(\hat{\Gamma}_{nm}c_0/\omega) = O(1)$ if $He < \gamma_{nm}$ and $\gamma_{nm} - He = O(1)$, $\Re(\hat{\Gamma}_{nm}c_0/\omega) = O(St^{-1/2})$ if $|He - \gamma_{nm}| \ll 1$, and $\Re(\hat{\Gamma}_{nm}c_0/\omega) = O(St^{-1})$ if $He > \gamma_{nm}$ and $He - \gamma_{nm} = O(1)$. In addition, Fig. 3 exemplarily compares the attenuation rate for the mode $(0, 1)$ in the region $He > \gamma_{01}$ with the different solutions that can be derived from the other theories mentioned before. It well illustrates the fact that for a propagating mode, the damping mechanisms in the core region make an important contribution to the correction of Beatty's solution⁷ and, therefore, cannot be neglected if higher order effects are to be considered.

B. The case $l \gg 1$

Inspection of the expression for the propagation parameters (80) shows that if the scaling parameter $l \sim He$ is assumed to be so large that $\alpha l^2 = O(1)$, the boundary layer effects and the viscothermal effects in the core region become of the same order of magnitude. In order to investigate this case in an asymptotically correct way and to maintain the linearity of the problem, l has to be scaled as

$$l = \alpha^{-1/2} l_1, \quad l_1 = O(1), \quad (84)$$

and M has to be chosen such that

$$M \ll \alpha. \quad (85)$$

Consequently, $Re^{-1} \sim St^{-1} \sim \alpha \ll 1$ and the conditions (2) and (5) are satisfied.

By performing a perturbation analysis very similar to that before, however, with $\alpha^{1/2}$ rather than α used as perturbation parameter, one obtains the following wave equations for the first-, second-, and third-order quantities:

$$\omega^2 \hat{p}_1 + \frac{\partial^2 \hat{p}_1}{\partial z^2} = 0, \quad (86)$$

$$\omega^2 \hat{p}_{\alpha^{1/2}} + \frac{\partial^2 \hat{p}_{\alpha^{1/2}}}{\partial z^2} = 0, \quad (87)$$

$$\omega^2 \hat{p}_{\alpha} + \frac{\partial^2 \hat{p}_{\alpha}}{\partial z^2} = -\frac{1}{l_1^2 r} \frac{\partial}{\partial r} \left(r \frac{\partial \hat{p}_1}{\partial r} \right) - \frac{1}{l_1^2 r^2} \frac{\partial \hat{p}_1}{\partial \phi^2} - 2 \frac{\partial^2 \hat{p}_1}{\partial z \partial z_1} + l_1^2 \frac{j \omega^3}{d^2} \hat{p}_1 \left(\frac{4}{3} + \eta_0 + \frac{\gamma_0 - 1}{Pr} \right). \quad (88)$$

Furthermore, the matching rules yield

$$r = 1: \quad \frac{\partial \hat{p}_1}{\partial r} = l_1^2 \omega^2 \hat{F} \hat{p}_1, \quad \hat{F} = \frac{1-j}{d \sqrt{2} \omega} \left(1 + \frac{\gamma_0 - 1}{\sqrt{Pr}} \right). \quad (89)$$

Obviously, the terms on the right-hand side of Eq. (88) are resonant and have to be annihilated since they would inevitably lead to secular solutions for \hat{p}_{α} . This generates a partial differential equation which, together with the boundary condition (89), determines the evolution of the leading order pressure fluctuations over distances of the order $O(\alpha^{-1})$:

$$2 \frac{\partial^2 \hat{p}_1}{\partial z \partial z_1} + \frac{1}{l_1^2 r} \frac{\partial}{\partial r} \left(r \frac{\partial \hat{p}_1}{\partial r} \right) + \frac{1}{l_1^2 r^2} \frac{\partial \hat{p}_1}{\partial \phi^2} - l_1^2 \frac{j \omega^3}{d^2} \hat{p}_1 \left(\frac{4}{3} + \eta_0 + \frac{\gamma_0 - 1}{Pr} \right) = 0. \quad (90)$$

Worth mentioning is the fact that Eqs. (86) and (90) constitute the linearized version of the Khokhlov-Zabolotskaya-Kuznetsov (KZK) model the nonlinear study by Makarov and Vatrushina¹⁴ is based on. Proceeding from the same assumptions concerning the thickness of the boundary layer and the ratio $\tilde{\lambda}/\tilde{R}$ as made here, they derived a boundary condition which is formally almost identical to Eq. (89). However, while adopting a general equation of state (for a real gas), the energy equation for a perfect gas was used and, thus, the result given in Ref. 14, Eq. (36), is not completely correct: The term $\beta c_0^2 / (Pr^{1/2} c_p)$ in this equation should read $T_0 \beta^2 c_0^2 / (Pr^{1/2} c_p)$, which, expressed in the notation used here, then simplifies to $G_0 \tilde{\delta}_0 \tilde{\beta}_0 / Pr^{1/2} = (\gamma_0 - 1) / Pr^{1/2}$.

The system (86), (89), and (90) can be solved to give

$$\hat{p}_1 = \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} \left(\hat{C}_{1nm1} \frac{\hat{J}_m(\hat{\gamma}_{nm} r)}{\hat{J}_m(\hat{\gamma}_{nm})} e^{-j[\omega(z + \hat{L}_{nm} z_1) + m\phi]} + \hat{C}_{2nm1} \frac{\hat{J}_m(\hat{\gamma}_{nm} r)}{\hat{J}_m(\hat{\gamma}_{nm})} e^{j[\omega(z + \hat{L}_{nm} z_1) + m\phi]} \right), \quad (91)$$

where the coefficients $\hat{\gamma}_{nm} = \hat{\gamma}_{n(-m)}$, $n = 0, 1, 2, \dots$, are the complex roots of

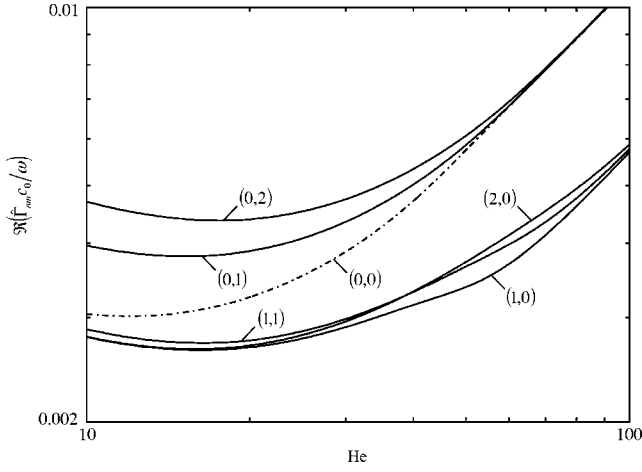


FIG. 4. Graphs of $\Re(\hat{\Gamma}_{nm}c_0/\omega)$ as function of $\text{He} \sim \text{St}^{1/2} \sim \text{Re}^{1/3} \gg 1$ in double logarithmic scale, $R=0.001$ m, i.e., $\text{St} = \sqrt{\text{He}} \text{Re}_R$ with $\text{Re}_R = c_0 R \rho_0 / \mu_0$; $\text{Pr}=0.707$, $\gamma_0=1.402$, $\eta_0/\mu_0=0.6$, $c_0 \rho_0 / \mu_0 = 221.3732 \times 10^5 \text{ m}^{-1}$ (air at 300 K and 10^5 Pa).

$$\hat{\mathbf{J}}_{m+1}(\hat{\gamma}_{nm})\hat{\gamma}_{nm} - m\hat{\mathbf{J}}_m(\hat{\gamma}_{nm}) + l_1^2 \omega^2 \hat{\mathbf{F}}\hat{\mathbf{J}}_m(\hat{\gamma}_{nm}) = 0 \quad (92)$$

and

$$\hat{L}_{nm} = -\frac{\hat{\gamma}_{nm}^2}{2l_1^2 \omega^2} - l_1^2 \frac{j\omega}{2d^2} \left(\frac{4}{3} + \eta_0 + \frac{\gamma_0 - 1}{\text{Pr}} \right). \quad (93)$$

This reveals the fact that in the case $l \sim \text{St}^{1/2}$ even the lowest mode $(n,m)=(0,0)$ does not propagate as a plane wave since all $\hat{\gamma}_{n0} \neq 0$.

Using the relationship (79), the propagation parameters turn out to be given by the expansion

$$\hat{\Gamma}_{nm} = j\omega + \alpha j\omega \hat{L}_{nm} + \dots \quad (94)$$

The corresponding expressions in terms of dimensional quantities are presented in Appendix A. As can be seen from comparison of the graphs of $\Re(\hat{\Gamma}_{nm}c_0/\omega)$ depicted in Fig. 4 with the results plotted in Figs. 2 and 3(a), the solutions from Eqs. (94) and (A3) indeed continue the solutions derived for the case $l \sim \text{He} = O(1)$ to the region $l \sim \text{He} = O(\text{St}^{1/2})$. Furthermore, one can immediately observe from Fig. 4 the growing influence of the viscothermal losses in the core region on the damping rates as the Helmholtz number increases.

IV. THE NONLINEAR CASE

In the following, the propagation and modulation of a finite-amplitude wave packet centered at the frequency ω_c and the wave number \hat{k}_c in a long hard-walled cylindrical duct shall be considered. As in Sec. III, the characteristic wavelength $\tilde{\lambda}$ is presupposed to be comparable to the radius of the tube \tilde{R} . Therefore, the scaling parameter $l \sim \text{He}_c \sim 1$ can be set to unity without loss of generality, resulting in $\omega_c = \text{He}_c$. The fluid is assumed to be an initially quiescent real gas featuring a uniform equilibrium state which satisfies the conditions (3). Furthermore, the analysis proceeds from the assumption that the spatial range consumed by the viscothermal boundary layer is so small that

$$\text{St}^{-1} \sim M^2, \quad \text{i.e., } \alpha = \sigma M^2, \quad (95)$$

holds, where $\sigma = O(1)$ or smaller. As shown by Nayfeh¹⁸ for the case of a perfect gas, then the evolution of the leading order pressure fluctuations are governed by an equation where boundary layer dissipation might become of the same order of magnitude as the nonlinear effects. For the level of approximation to be reached in this study, the thermoviscous losses in the core region will turn out to be negligibly small since these effects are an order of magnitude smaller than those of the wall layer.

According to Eq. (95), the perturbation analysis can be carried out with respect to the remaining parameter M only. However, the determination of an asymptotically correct solution, incorporating the effects of the wave modulation with time and space as well as the nonlinearity, necessitates the introduction of the following spatial and temporal scales:

$$\frac{\partial}{\partial z} \rightarrow \frac{\partial}{\partial z} + M \frac{\partial}{\partial z_1} + M^2 \frac{\partial}{\partial z_2}, \quad \frac{\partial}{\partial t} \rightarrow \frac{\partial}{\partial t} + M \frac{\partial}{\partial t_1} + M^2 \frac{\partial}{\partial t_2}. \quad (96)$$

In this connection it should be noted that in the linear case discussed in Sec. III, the use of additional temporal scales $t_1 = \alpha t$ and $t_2 = \alpha^2 t$ would lead to solutions which can be written in the form [cf. Eqs. (23) and (24)]

$$w(z, z_1, z_2, r, \phi, t, t_1, t_2) = \Re(\hat{w}_{st}(z, z_1, z_2, r, \phi) e^{j(\omega t + \omega_\alpha t_1 + \omega_{\alpha^2} t_2)}), \quad (97)$$

where $\omega_\alpha = O(1)$ and $\omega_{\alpha^2} = O(1)$ denote arbitrary frequency detunings. Consequently, they can again be decomposed into steady amplitude function \hat{w}_{st} and time-harmonic functions. The same results would be obtained, however, by letting $\omega \rightarrow \omega + \alpha \omega_\alpha + \alpha^2 \omega_{\alpha^2}$ and expanding the parameters \hat{k}_{nm} , \hat{F}_{nm} and \hat{H}_{nm} appearing in the solutions (74) and (75) with respect to α . As demonstrated in Sec. III A, this method can even be applied to calculate the linear propagation parameters for driving frequencies close to the cutoff frequency of a particular mode, where the attenuation of this mode takes place over spatial ranges $z = O(\alpha^{-1/2})$.

A further consequence of the condition (95) is that in the core region where $r = O(1)$ the right-hand side of Eq. (13) is of the order $O(M^5)$. Utilizing the asymptotic expansions for ρ , ϑ , and s , the energy equation thus reduces to

$$\frac{ds_1}{dt} = \frac{ds_M}{dt} = \frac{ds_{M^2}}{dt} = 0. \quad (98)$$

Hence, after expanding the total derivative

$$(\delta p) = \left(\frac{\partial p}{\partial \rho} \right)_s (\delta \rho) + \left(\frac{\partial p}{\partial s} \right)_\rho (\delta s) \quad (99)$$

with respect to M and substituting the parameters Γ_0 and Λ_0 defined in Eqs. (1) and (3), integration yields

$$\begin{aligned} \rho_1 &= p_1, & \rho_M &= p_M - (\Gamma_0 - 1)p_1^2, \\ \rho_{M^2} &= p_{M^2} - 2(\Gamma_0 - 1)p_1 p_M - \frac{1}{3}(\Lambda_0 - 4\Gamma_0^2 + 7\Gamma_0 - 3)p_1^3. \end{aligned} \quad (100)$$

Furthermore, evaluation of the Navier-Stokes equations (9)–(11) for the leading order perturbations in the core region implies that

$$\frac{\partial}{\partial t}(\nabla \times \mathbf{v}_1) = \mathbf{0} \Rightarrow \nabla \times \mathbf{v}_1 = \mathbf{0}, \quad (101)$$

in which

$$\nabla \times \mathbf{A} = \left(\frac{1}{r} \left(\frac{\partial(rA_\phi)}{\partial r} - \frac{\partial A_r}{\partial \phi} \right), \frac{1}{r} \frac{\partial A_z}{\partial \phi} - \frac{\partial A_\phi}{\partial z}, \frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r} \right)^T \quad (102)$$

denotes the vorticity of an arbitrary vector \mathbf{A} . This enables to define a velocity potential Φ_1 ,

$$\mathbf{v}_1 = \nabla \Phi_1, \quad (103)$$

whereupon substitution for ρ_1 and \mathbf{v}_1 from Eqs. (100) and (103) into the leading order outer expansions of the Navier-Stokes equations and the continuity equation (12) gives

$$p_1 = -\frac{\partial \Phi_1}{\partial t}, \quad \frac{\partial^2 \Phi_1}{\partial t^2} - \Delta \Phi_1 = 0. \quad (104)$$

The equations of motion (9)–(11) together with Eq. (98) show that for all expansion terms at least up to the order $O(M^3)$, the flow in the core region can be regarded as isentropic and inviscid. By invoking Helmholtz's theorems of vorticity, one thus can expect Eq. (101) to hold for \mathbf{v}_M and \mathbf{v}_{M^2} as well. Indeed, evaluation of the outer expansions of Eqs. (9)–(11) for the second-order terms immediately verifies this assumption in the generalized sense that now

$$\nabla \times \left(\mathbf{v}_M - \left(\frac{\partial \Phi_1}{\partial z_1}, 0, 0 \right)^T \right) = \mathbf{0} \Rightarrow \mathbf{v}_M = \nabla \Phi_M + \left(\frac{\partial \Phi_1}{\partial z_1}, 0, 0 \right)^T. \quad (105)$$

Combining this result with Eq. (100) and the second-order expansion of the continuity equation (12), then the set

$$p_M = -\frac{\partial \Phi_M}{\partial t} - \frac{\partial \Phi_1}{\partial t_1} - \frac{1}{2} \left[(\nabla \Phi_1)^2 - \left(\frac{\partial \Phi_1}{\partial t} \right)^2 \right], \quad (106)$$

$$\frac{\partial^2 \Phi_M}{\partial t^2} - \Delta \Phi_M = 2 \left(\frac{\partial^2 \Phi_1}{\partial z \partial z_1} - \frac{\partial^2 \Phi_1}{\partial t \partial t_1} \right) - 2(\Gamma_0 - 1) \Delta \Phi_1 \frac{\partial \Phi_1}{\partial t} - \frac{\partial}{\partial t} (\nabla \Phi_1)^2.$$

is obtained.

Proceeding in very much the same way as before, the expansions of the Navier-Stokes equations governing the third-order perturbations in the core region show that

$$\nabla \times \left(\mathbf{v}_{M^2} - \left(\frac{\partial \Phi_1}{\partial z_2} + \frac{\partial \Phi_M}{\partial z_1}, 0, 0 \right)^T \right) = \mathbf{0} \Rightarrow \mathbf{v}_{M^2} = \nabla \Phi_{M^2} + \left(\frac{\partial \Phi_1}{\partial z_2} + \frac{\partial \Phi_M}{\partial z_1}, 0, 0 \right)^T. \quad (107)$$

This leads to an expression for p_{M^2} in terms of the potentials, which together with the expansion terms for ρ from Eq. (100) is substituted into the third-order outer expansion of Eq. (12). Thus, the resulting equations read

$$\begin{aligned} p_{M^2} = & -\frac{\partial \Phi_{M^2}}{\partial t} - \frac{\partial \Phi_M}{\partial t_1} - \frac{\partial \Phi_1}{\partial t_2} + \frac{\partial \Phi_M}{\partial t} \frac{\partial \Phi_1}{\partial t} - \nabla \Phi_M \cdot \nabla \Phi_1 \\ & + \frac{\partial \Phi_1}{\partial t} \frac{\partial \Phi_1}{\partial t_1} - \frac{\partial \Phi_1}{\partial z} \frac{\partial \Phi_1}{\partial z_1} + \frac{1}{6} \frac{\partial \Phi_1}{\partial t} \left[2\Gamma_0 \left(\frac{\partial \Phi_1}{\partial t} \right)^2 \right. \\ & \left. - 3 \left(\frac{\partial \Phi_1}{\partial t} \right)^2 + 3(\nabla \Phi_1)^2 \right], \\ & \frac{\partial^2 \Phi_{M^2}}{\partial t^2} - \Delta \Phi_{M^2} \\ = & 2 \left(\frac{\partial^2 \Phi_M}{\partial z \partial z_1} - \frac{\partial^2 \Phi_M}{\partial t \partial t_1} + \frac{\partial^2 \Phi_1}{\partial z \partial z_2} - \frac{\partial^2 \Phi_1}{\partial t \partial t_2} \right) + \frac{\partial^2 \Phi_1}{\partial z_1^2} - \frac{\partial^2 \Phi_1}{\partial t_1^2} \\ & - 2(\Gamma_0 - 1) \left[\left(\Delta \Phi_M + 2 \frac{\partial^2 \Phi_1}{\partial z \partial z_1} \right) \frac{\partial \Phi_1}{\partial t} \right. \\ & \left. + \Delta \Phi_1 \left(\frac{\partial \Phi_M}{\partial t} + \frac{\partial \Phi_1}{\partial t_1} \right) \right] \\ & - 2 \frac{\partial}{\partial t} \left(\nabla \Phi_M \cdot \nabla \Phi_1 + \frac{\partial \Phi_1}{\partial z} \frac{\partial \Phi_1}{\partial z_1} \right) - \frac{\partial}{\partial t_1} (\nabla \Phi_1)^2 \\ & + \Lambda_0 \Delta \Phi_1 \left(\frac{\partial \Phi_1}{\partial t} \right)^2 - (\Gamma_0 - 1) \Delta \Phi_1 (\nabla \Phi_1)^2 \\ & - \frac{1}{2} \nabla \Phi_1 \cdot \nabla (\nabla \Phi_1)^2. \end{aligned} \quad (108)$$

In order to find the boundary conditions for the three wave equations governing the potential functions Φ_1 , Φ_M , and Φ_{M^2} , respectively, it is necessary to investigate the acoustic motion in the wall layer. Since the scaling parameter α is presupposed to be of the order $O(M^2)$, the equations resulting from the inner expansion will turn out to be linear, suggesting the application of a complex Fourier transform

$$\hat{F}[A](\omega) = \int_{-\infty}^{\infty} A(t) e^{-j\omega t} dt \quad (109)$$

with respect to the ‘‘fast’’ time t . Thus, most of the results already derived in Sec. III for the inner radial velocity terms up to the order $O(\alpha)$ carry over almost unchanged to the case considered here: As a consequence of $\hat{F}[V_{r1}] = 0$ and, in extension to Eq. (27), $\hat{F}[V_{rM}] = 0$, Van Dyke's matching conditions imply that

$$r = 1: \quad \frac{\partial \Phi_1}{\partial r} = 0, \quad \frac{\partial \Phi_M}{\partial r} = 0. \quad (110)$$

In addition, by setting $\alpha = \sigma M^2$, $l = 1$, and applying the matching rules to the inner and outer leading order terms of the axial and azimuthal velocities and the density, Eq. (53) can be recast into

$$\begin{aligned} \hat{F}[V_{z1}] &= \hat{F} \left[\frac{\partial \Phi_1}{\partial z} \right]_{r=1} (1 - e^{-(1+j)(d/\sigma)y\sqrt{\omega/2} \operatorname{sgn}(\omega)}), \\ \hat{F}[V_{\phi 1}] &= \hat{F} \left[\frac{\partial \Phi_1}{\partial \phi} \right]_{r=1} (1 - e^{-(1+j)(d/\sigma)y\sqrt{\omega/2} \operatorname{sgn}(\omega)}), \end{aligned} \quad (111)$$

$$\hat{F}[\Psi_1] = -j\omega\hat{F}[\Phi_1]_{r=1} [1 + (\gamma_0 - 1) \times (1 - e^{-(1+j)(d/\sigma)y\sqrt{\omega\text{Pr}/2} \text{sgn}(\omega)})].$$

The solution for $\hat{F}[V_{rM^2}]$ corresponding to Eq. (54) then assumes the form

$$\begin{aligned} \hat{F}[V_{rM^2}] = & -j\omega\hat{F}[\Phi_1]_{r=1} \left[j\omega y + \sigma \frac{1+j}{d} \sqrt{\frac{\omega}{2}} \text{sgn}(\omega) \right. \\ & \times \frac{\gamma_0 - 1}{\sqrt{\text{Pr}}} (1 - e^{-(1+j)(d/\sigma)y\sqrt{\omega\text{Pr}/2} \text{sgn}(\omega)}) \left. \right] \\ & - \frac{j}{\omega} \hat{F} \left[\frac{\partial^2 \Phi_1}{\partial z^2} + \frac{\partial^2 \Phi_1}{\partial \phi^2} \right]_{r=1} \left[j\omega y - \sigma \frac{1+j}{d} \sqrt{\frac{\omega}{2}} \right. \\ & \times \text{sgn}(\omega) (1 - e^{-(1+j)(d/\sigma)y\sqrt{\omega/2} \text{sgn}(\omega)}) \left. \right], \quad (112) \end{aligned}$$

and the matching principle applied to the expansion terms of v_r and V_r yields the boundary condition

$$\begin{aligned} r=1: \quad \hat{F} \left[\frac{\partial \Phi_{M^2}}{\partial r} \right] = & \sigma \frac{1-j}{d} \omega \sqrt{\frac{\omega}{2}} \text{sgn}(\omega) \\ & \times \left[\hat{F}[\Phi_1] \frac{\gamma_0 - 1}{\sqrt{\text{Pr}}} - \left(\hat{F} \left[\frac{\partial^2 \Phi_1}{\partial z^2} \right] + \hat{F} \left[\frac{\partial^2 \Phi_1}{\partial \phi^2} \right] \right) \frac{1}{\omega^2} \right]. \quad (113) \end{aligned}$$

The goal pursued in the study presented here is to derive the evolution equation for the spatial and temporal modulation of a traveling wave packet centered at the frequency $\omega_c > 0$ and the real wave number $k_c > 0$. Therefore, the solution to Eq. (104) subject to the boundary condition for Φ_1 from Eq. (110) is taken to be the right-running wave

$$\begin{aligned} \Phi_1 = \hat{\Phi}_1 + \hat{\Phi}_1^* = & \hat{A}_1(z_1, z_2, t_1, t_2) \frac{J_{m_c}(\gamma_c r)}{J_{m_c}(\gamma_c)} e^{j\zeta} + \text{CC}, \\ \zeta = \omega_c t - k_c z - m_c \phi, \quad (114) \end{aligned}$$

where the superscript “*” denotes the conjugate complex of a quantity, CC stands for the conjugate complex of the preceding terms, \hat{A}_1 is the weakly varying shape function of the carrier mode $(n, m) = (n_c, m_c)$, $\gamma_c = \gamma_{n_c, m_c}$ is a zero of $\partial J_{m_c}(\xi)/\partial \xi$, and k_c necessarily satisfies the leading order dispersion relation (51) for a propagating mode with $\omega_c > \gamma_c > 0$ ($l=1$). As outlined in Sec. II A, the case of the carrier mode being a plane wave, i.e., $m_c = \gamma_c = 0$, has to be excluded since otherwise the interaction with all its harmonics would have to be considered as well. Hence, by taking advantage of the solution (114), the boundary condition (113) for the wave equation (108) is modified to

$$r=1: \quad \frac{\partial \Phi_{M^2}}{\partial r} = \sigma \hat{\Phi}_1 \frac{1-j}{d\sqrt{2\omega_c}} \left(\omega_c^2 \frac{\gamma_0 - 1}{\sqrt{\text{Pr}}} + k_c^2 + m_c^2 \right) + \text{CC}. \quad (115)$$

When evaluated for a perfect gas with constant specific heats, i.e., $\Gamma_0 = (\gamma + 1)/2$, $\gamma = \text{const}$, and $\Lambda_0 = 0$, the set of equations (104), (106), and (108) immediately turns out to be

in complete conformity with the wave equations that form the basis for Nayfeh’s investigations.¹⁸ Following the line of reasoning given in Ref. 18, the derivations of the evolution equations determining the behavior of the amplitude \hat{A}_1 over the long spatial and temporal ranges z_1, z_2, t_1 , and t_2 can thus be developed in a straightforward way invoking the solvability conditions provided by the *Fredholm alternative* in the second- and third-order problems. This concept is also known as the so-called *reductive perturbation method*.³²

The solutions for the second-order problem, i.e., Eq. (106) subject to condition (110), can be summarized as follows:

$$k_c \frac{\partial \hat{A}_1}{\partial z_1} + \omega_c \frac{\partial \hat{A}_1}{\partial t_1} = 0, \quad (116)$$

$$\begin{aligned} \Phi_M = & \frac{\hat{A}_1^2}{J_{m_c}^2(\gamma_c)} (\hat{C}_1 J_{m_c}^2(\gamma_c r) + \hat{C}_2 r J_{m_c}(\gamma_c r) J_{m_c+1}(\gamma_c r) \\ & + \hat{C}_3 J_{2m_c}(2\gamma_c r)) + \text{CC}, \\ \hat{C}_1 = & j\omega_c \left(\Gamma_0 \frac{\omega_c^2}{\gamma_c^2} m_c + 1 \right), \quad \hat{C}_2 = -j\Gamma_0 \frac{\omega_c^3}{\gamma_c}, \quad (117) \end{aligned}$$

$$\hat{C}_3 = -\hat{C}_2 \frac{J_{m_c}(\gamma_c)}{2\gamma_c J'_{2m_c}(\gamma_c)} (J_{m_c+1}(\gamma_c) + \gamma_c J'_{m_c+1}(\gamma_c)).$$

Please note that the expression for \hat{C}_3 printed in Ref. 18 [Eq. (39b)] is incorrect and had to be recalculated.

The evolution equation governing the propagation of the leading order solution Φ_1 over distances and periods of the order $O(M^{-2})$ can be determined in a very similar manner, that is, by substituting the expressions for Φ_1 and Φ_M from Eqs. (114) and (117) into the third-order wave equation (108) and calculating the solvability condition for the forcing terms proportional to $\exp(j\zeta)$. In contrast to the second-order problem, however, the application of the Fredholm alternative [see Eq. (35) in Ref. 18] leads to an integral condition which has to be evaluated using the inhomogeneous boundary condition (115) at $r=1$. The resulting equation then reads

$$\begin{aligned} 2j \left(k_c \frac{\partial \hat{A}_1}{\partial z_2} + \omega_c \frac{\partial \hat{A}_1}{\partial t_2} \right) - \frac{\partial^2 \hat{A}_1}{\partial z_1^2} + \frac{\partial^2 \hat{A}_1}{\partial t_1^2} \\ = -2N \hat{A}_1^2 \hat{A}_1^* + 2\sigma k_c^2 \hat{F}_c \hat{A}_1, \quad (118) \end{aligned}$$

where $\hat{F}_c = \hat{F}_{n_c, m_c}$ is the quantity given by Eq. (58) and the nonlinearity parameter N is defined as

$$N = \frac{1}{J_{m_c}^4(\gamma_c) (1 - m_c^2/\gamma_c^2)} \int_0^1 r M(r) J_{m_c}(\gamma_c r) dr, \quad (119)$$

in which

$$\begin{aligned}
M(r) = & \gamma_c^2 J_{m_c}(\gamma_c r) J_{m_c}'^2(\gamma_c r) \left[(\Gamma_0 - 1) \omega_c^2 + \frac{m_c(5m_c - 3)}{r^2} - 3\gamma_c^2 + 2k_c^2 \right] - \gamma_c J_{m_c}^2(\gamma_c r) J_{m_c}'(\gamma_c r) \left[\frac{m_c^2}{r^3} - 2j\omega_c \left(\frac{2\hat{C}_1 m_c}{r} + \hat{C}_2 \gamma_c r \right) \right] \\
& + \frac{3\gamma_c^3}{r} J_{m_c}'^2(\gamma_c r) J_{m_c+1}(\gamma_c r) + 2j\omega_c \gamma_c J_{m_c}'(\gamma_c r) \left(-\hat{C}_2 \gamma_c r J_{m_c+1}^2(\gamma_c r) - 2\hat{C}_1 \gamma_c J_{m_c}(\gamma_c r) J_{m_c+1}(\gamma_c r) \right. \\
& \left. + \frac{2\hat{C}_3 m_c}{r} J_{2m_c}(2\gamma_c r) - 2\hat{C}_3 \gamma_c J_{2m_c+1}(2\gamma_c r) \right) - \left[2(\Gamma_0 - 1) \omega_c^2 + \frac{2m_c^2}{r^2} + 2k_c^2 \right] \\
& \times \left[J_{m_c}^3(\gamma_c r) \left(\frac{m_c^2}{2r^2} + \frac{k_c^2}{2} + 2(\Gamma_0 - 1) \omega_c^2 - 2j\hat{C}_1 \omega_c \right) - 2j\omega_c J_{m_c}(\gamma_c r) (\hat{C}_2 r J_{m_c}(\gamma_c r) J_{m_c+1}(\gamma_c r) + \hat{C}_3 J_{2m_c}(2\gamma_c r)) \right] \\
& + \Lambda_0 \omega_c^4 J_{m_c}^3(\gamma_c r). \tag{120}
\end{aligned}$$

Surprisingly, with the expressions for \hat{C}_1 , \hat{C}_2 , and \hat{C}_3 from Eq. (117) substituted into Eq. (120), all imaginary terms in $M(r)$ cancel out, leading to N being a purely real quantity. Furthermore, when Γ_0 is formally replaced with $(\gamma + 1)/2$, Λ_0 is set to zero, and j is replaced with $-i$, one obtains precisely the function $F(r)$ derived in Ref. 18 for the case of a perfect gas. In this connection it should be mentioned that the terms $3\kappa^2/r J_m'^2(\kappa r) J_{m+1}(\kappa r)$ and $\Gamma_2 \gamma J_m(\kappa r) J_{m+1}(\kappa r)$ appearing in the definition of $F(r)$ should read $3\kappa^3/r J_m'^2(\kappa r) J_{m+1}(\kappa r)$ and $\Gamma_2 r J_m(\kappa r) J_{m+1}(\kappa r)$, respectively. Moreover, upon neglecting the nonlinear and the temporal modulation terms in Eqs. (116) and (118), the solvability condition (59) from Sec. III for a linear monochromatic wave is recovered.

By exploiting Eq. (116) to eliminate the derivative $\partial^2 \hat{A}_1 / \partial z_1^2$ appearing in Eq. (118), it follows that

$$\frac{\partial \hat{A}_1}{\partial z_2} + \frac{1}{\omega_g} \frac{\partial \hat{A}_1}{\partial t_2} + j \frac{\omega_g'}{2\omega_g^3} \frac{\partial^2 \hat{A}_1}{\partial t_1^2} = j \frac{N}{k_c} \hat{A}_1^2 \hat{A}_1^* - j \sigma k_c \hat{F}_c \hat{A}_1, \tag{121}$$

where

$$\begin{aligned}
\omega_g &= \left(\frac{\partial \omega}{\partial \hat{k}_{nm}} \right)_c = \left(\frac{\partial}{\partial \hat{k}_{nm}} \sqrt{\hat{k}_{nm}^2 + \gamma_c^2} \right) \Big|_{\hat{k}_{nm}=k_c} = \frac{k_c}{\omega_c} > 0, \\
\omega_g' &= \left(\frac{\partial^2 \omega}{\partial \hat{k}_{nm}^2} \right)_c = \frac{1}{\omega_c} \left(1 - \frac{k_c^2}{\omega_c^2} \right) > 0
\end{aligned} \tag{122}$$

are, respectively, the group velocity and its first derivative with respect to the axial wave number. Furthermore, adding Eq. (116) multiplied with M/k_c to Eq. (121) multiplied with M^2 , reintroducing the original temporal and spatial coordinates t and z , and neglecting terms of the order $O(M^3)$ (which is asymptotically correct) results in

$$\begin{aligned}
\frac{\partial \hat{A}_1}{\partial z} + \frac{1}{\omega_g} \frac{\partial \hat{A}_1}{\partial t} + j \frac{\omega_g'}{2\omega_g^3} \frac{\partial^2 \hat{A}_1}{\partial t^2} \\
= j M^2 \frac{N}{k_c} \hat{A}_1^2 \hat{A}_1^* - j M^2 \sigma k_c \hat{F}_c \hat{A}_1. \tag{123}
\end{aligned}$$

To simplify the following discussions, the suitably scaled quantities \hat{U} , D , X and the wave coordinate ξ , with

$$Z = M^2 z, \quad \xi = M \sqrt{\frac{\omega_g}{\omega_g'}} (\omega_g t - z), \tag{124}$$

$$D = -k_c \Im(\hat{F}_c) = k_c \Re(\hat{F}_c) > 0, \quad \hat{U} = \sqrt{\frac{|N|}{k_c}} \hat{A}_1^* e^{-j\sigma D Z},$$

are introduced, and one arrives at a linearly damped *nonlinear Schrödinger equation* (NLS) in normalized form:

$$j \frac{\partial \hat{U}}{\partial Z} + \frac{1}{2} \frac{\partial^2 \hat{U}}{\partial \xi^2} - \text{sgn}(N) \hat{U}^2 \hat{U}^* = -j \sigma D \hat{U}, \tag{125}$$

where the term involving the second derivative represents the dispersive effects induced by the temporal modulation of the wave packet, the third term on the left-hand side describes the nonlinearity of the system, and the term proportional to σ results from the dissipative effects of the viscothermal boundary layer.

A. Comments on the damping parameter

For Eq. (125) to provide an asymptotically correct description of the most relevant physical mechanisms present in the problem, in principle, the damping parameter σ has to be assumed to be an order $O(1)$ quantity; that is to say, a balance between dispersion, nonlinearity, and dissipation has to exist. However, if the boundary layer effects are small but not so small that additional nonlinear effects of higher order come into play as well, i.e., if $M \ll \sigma \ll 1$ or, equivalently, $M^6 \ll \text{Re}^{-1} \sim \alpha^2 \ll M^4$ [see Eqs. (8) and (95)], one may try to expand the scaled shape function \hat{U} with respect to the damping parameter and introduce an additional “long” length scale Z_σ which will account for the slow variations in longitudinal direction:

$$\hat{U} = \hat{U}_1 + \sigma \hat{U}_\sigma + \dots, \quad \frac{\partial}{\partial Z} \rightarrow \frac{\partial}{\partial Z} + \sigma \frac{\partial}{\partial Z_\sigma}. \tag{126}$$

These expressions substituted into Eq. (125) then generate the system

$$M \ll \sigma \ll 1: \quad j \frac{\partial \hat{U}_1}{\partial Z} + \frac{1}{2} \frac{\partial^2 \hat{U}_1}{\partial \xi^2} - \text{sgn}(N) \hat{U}_1^2 \hat{U}_1^* = 0, \tag{127}$$

$$j \frac{\partial \hat{U}_\sigma}{\partial Z} + \frac{1}{2} \frac{\partial^2 \hat{U}_\sigma}{\partial \xi^2} - \text{sgn}(N) \hat{U}_1 (2 \hat{U}_1^* \hat{U}_\sigma + \hat{U}_1 \hat{U}_\sigma^*)$$

$$= -j \frac{\partial \hat{U}_1}{\partial Z_\sigma} - j D \hat{U}_1. \quad (128)$$

The same set of equations would have been obtained if the velocity potential Φ had been expanded with respect to the parameters M and σ from the beginning and, in addition to z_1, t_1, z_2, t_2 , also the spatial and temporal scales $z_\sigma = \sigma M^2 z, t_\sigma = \sigma M^2 t$ had been introduced. In other words, Eqs. (127) and (128) can be interpreted as constituting the solvability conditions for the problems at the orders $O(M^3)$ and $O(\sigma M^3)$, respectively. Consequently, Eq. (125) turns out to retain its validity even under conditions where dissipation is small, i.e., for all values of σ within the range $M \ll \sigma \leq O(1)$. If, however, $M \ll \sigma \ll 1$, the (linear) Eq. (128) can be used to derive an additional solvability condition that determines the variation of the leading order perturbation $\Phi_1 \propto \hat{U}_1$ over distances of the order $O(\sigma^{-1} M^{-2})$ and the changes of $\Phi_\sigma \propto \hat{U}_\sigma$ over ranges of the order $O(M^{-2})$.

B. Comments on the nonlinearity parameter

The integrand $rM(r)J_{m_c}(\gamma_c r)$ appearing in the definition (119) of the nonlinearity parameter N can be simplified significantly by expressing it in terms of the functions $J_{m_c}^4(\gamma_c r)r, J_{m_c}^2(\gamma_c r)J_{m_c+1}^2(\gamma_c r)r$ and the derivatives of $J_{m_c}^4(\gamma_c r)r^2, J_{m_c}^4(\gamma_c r), J_{m_c}^4(\gamma_c r)/r^2, J_{m_c}^3(\gamma_c r)J_{m_c}'(\gamma_c r)r, J_{m_c}^3(\gamma_c r)J_{m_c}'(\gamma_c r)/r, J_{m_c}^2(\gamma_c r)J_{m_c}''(\gamma_c r)r^2, J_{m_c}^2(\gamma_c r)J_{m_c}''(\gamma_c r), J_{m_c}(\gamma_c r)J_{m_c}'''(\gamma_c r)r, J_{m_c}'(\gamma_c r)r^2, J_{m_c}(\gamma_c r)J_{2m_c}(2\gamma_c r), J_{m_c}^2(\gamma_c r)J_{2m_c}(2\gamma_c r)r^2, J_{m_c}(\gamma_c r)J_{m_c}'(\gamma_c r)J_{2m_c}'(2\gamma_c r)r^2, J_{m_c}''(\gamma_c r)J_{2m_c}(2\gamma_c r)r^2, J_{m_c}^2(\gamma_c r)J_{2m_c+1}(2\gamma_c r)r, J_{m_c}^2(\gamma_c r)J_{2m_c+1}(2\gamma_c r)/r, J_{m_c}^2(\gamma_c r)J_{2m_c+1}'(2\gamma_c r)$. Integration with respect to r then yields

$$N = N_0 + \Gamma_0(\omega_c^2 N_1 + \omega_c^4 N_2) + \Gamma_0^2(\omega_c^4 N_3 + \omega_c^6 N_4) + \Lambda_0 \omega_c^4 N_5, \quad (129)$$

where

$$N_0 = -4\gamma_c^4 I_0 + \frac{3\gamma_c^4 - 8\gamma_c^4 m_c^2 + 8\gamma_c^2 m_c - 2\gamma_c^2 m_c^2 - 32\gamma_c^2 m_c^3 + 16\gamma_c^2 m_c^4 + 3m_c^4 - 8m_c^6}{4(1 - 4m_c^2)(1 - m_c^2/\gamma_c^2)},$$

$$N_1 = 8\gamma_c^2 I_0 - \frac{4m_c}{1 - m_c^2/\gamma_c^2}, \quad N_2 = -3I_1, \quad N_3 = -4I_1,$$

$$N_4 = \frac{1}{\gamma_c^2} \left(2I_1 - \frac{1}{1 - m_c^2/\gamma_c^2} \right) + \left(1 - \frac{m_c^2}{\gamma_c^2} \right) \frac{J_{2m_c}(2\gamma_c)}{\gamma_c J_{2m_c+1}(2\gamma_c) - m_c J_{2m_c}(2\gamma_c)}, \quad (130)$$

$$N_5 = I_1, \quad I_0 = \frac{1}{J_{m_c}^4(\gamma_c)(1 - m_c^2/\gamma_c^2)} \int_0^1 r J_{m_c}^2(\gamma_c r) (J_{m_c}^2(\gamma_c r) - J_{m_c+1}^2(\gamma_c r)) dr,$$

$$I_1 = \frac{1}{J_{m_c}^4(\gamma_c)(1 - m_c^2/\gamma_c^2)} \int_0^1 r J_{m_c}^4(\gamma_c r) dr.$$

The numerical effort for the evaluation of N thus reduces to the determination of the integrals I_0 and I_1 . Moreover, for the symmetrical carrier modes, the expressions from Eq. (130) simplify to

$$m_c = 0: \quad N_0 = -\frac{8\gamma_c^4 I_1}{3} + \frac{3\gamma_c^4}{4}, \quad N_1 = \frac{16\gamma_c^2 I_1}{3},$$

$$N_2 = -3I_1, \quad N_3 = -4I_1, \quad (131)$$

$$N_4 = \frac{2I_1 - 1}{\gamma_c^2} + \frac{J_0(2\gamma_c)}{\gamma_c J_1(2\gamma_c)}, \quad N_5 = I_1,$$

$$I_1 = \frac{1}{J_0^4(\gamma_c)} \int_0^1 r J_0^4(\gamma_c r) dr.$$

In this case, $I_0 = 2I_1/3$ and, therefore, only the computation of I_1 is necessary.

The values of γ_c, I_0 , and I_1 calculated for various radial and azimuthal modes are presented in Table I, which enables convenient evaluation of the parameter N for different frequencies. For example, by setting $\Gamma_0 = 1.20$ and $\Lambda_0 = 0$ (air as perfect gas), the results given by the curves in Figs. 1 and 2 of Ref. 18 are reobtained. Please note that in that study, the nonlinearity parameter is defined as $\Lambda = 2NJ_0^2(\gamma_c)$. Furthermore, an error in the labeling of the mentioned figures has to be corrected here: the ordinates should be labeled $1/2\Lambda k^{-1}$ rather than $1/8\Lambda k^{-1}$. As far as the symmetrical modes are concerned, the solution for N from Eqs. (129) and (131) can also be compared with the results for the nonlinear wave number shift in a perfect gas derived by Keller and Millman¹⁶ using the method of strained parameters. Table I in Ref. 16 displays the numerical values for the coefficient $c^2 R^5 k_n^{(2)}$ at $\omega/c = 1.5\gamma_5$, which, according to the notation used here, equals $-N/(2\pi k_c)$ at

TABLE I. Numerical values of γ_c and the integrals I_0 and I_1 for different carrier modes (n_c, m_c) .

(n_c, m_c)	γ_c	I_0	I_1
(1, 0)	3.831 71		1.275 77
(2, 0)	7.015 59		1.481 02
(3, 0)	10.173 47		1.613 99
(4, 0)	13.323 69		1.712 35
(5, 0)	16.470 63		1.790 38
(0, 1)	1.841 18	0.337 63	0.411 77
(1, 1)	5.331 44	0.430 20	0.867 64
(2, 1)	8.536 32	0.535 94	1.043 43
(3, 1)	11.706 00	0.611 04	1.161 03
(0, 2)	3.054 24	0.282 19	0.390 84
(1, 2)	6.706 13	0.324 80	0.772 36
(2, 2)	9.969 47	0.410 29	0.927 93
(3, 2)	13.170 37	0.475 25	1.034 44
(0, 3)	4.201 19	0.250 53	0.380 74
(1, 3)	8.015 24	0.271 59	0.727 00
(2, 3)	11.345 92	0.344 23	0.869 74
(3, 3)	14.585 85	0.401 91	0.968 48

$\omega = 24.70594$. Precisely the same results are obtained when the parameter N is evaluated by utilizing Table I and setting $\Gamma_0 = 1.20$ and $\Lambda_0 = 0$.

As it will be discussed in the following, the type of (stable) solutions the damped NLS equation (125) has crucially depends on the sign of the nonlinearity parameter. For all modes considered in Table I, the following relationships hold: $N_0 < 0$, $N_1 > 0$, $N_4 > 0$, $N_5 = -1/3N_2 = -1/4N_3 > 0$. When the driving frequency becomes considerably large, $N \propto \Gamma_0^2 \omega_c^6 N_4$ and, consequently, the nonlinearity parameter will always be positive in the high-frequency range. However, in the low- and mid-frequency range, terms proportional to ω_c^2 and ω_c^4 come into play as well. In order for the latter to become negative, it is necessary and sufficient that $\Lambda_0 < 3\Gamma_0 + 4\Gamma_0^2$. Therefore, N is very likely to be negative within a finite band of frequencies if the gas parameter Λ_0 at the equilibrium reference state is sufficiently large and negative and Γ_0 is positive, but rather small such that the magnitude of the terms $\propto \omega_c^6$ is reduced.

For a wide range of fluids, such as perfect gases, the fundamental derivative is so large that N will turn out to be positive in the complete frequency range $\omega_c > \gamma_c > 0$. (In Ref. 18 this was shown for the case of air.) However, as mentioned in the Introduction, there exists a special class of real gases called BZT fluids which have a negative Γ region extending over a finite range of pressures and densities in the vicinity of the critical point. If such a fluid, therefore, expands along an isentrope crossing that region, starting from conditions with $\Gamma > 0$, the fundamental derivative changes its sign twice and will, eventually, become positive again. Obviously, at the second transition point where $\Gamma_0 = 0$ the parameter Λ_0 is negative and, hence, $N < 0$. Generally speaking, a BZT fluid always has equilibrium states where the nonlinearity parameter N is negative for a very wide range of driving frequencies.

To demonstrate this graphically, numerical results for N for the fluorocarbon FC-71^{23,24} are depicted in Fig. 5. The values for Γ_0 and Λ_0 at two different equilibrium reference states were calculated by applying the Martin-Hou equation

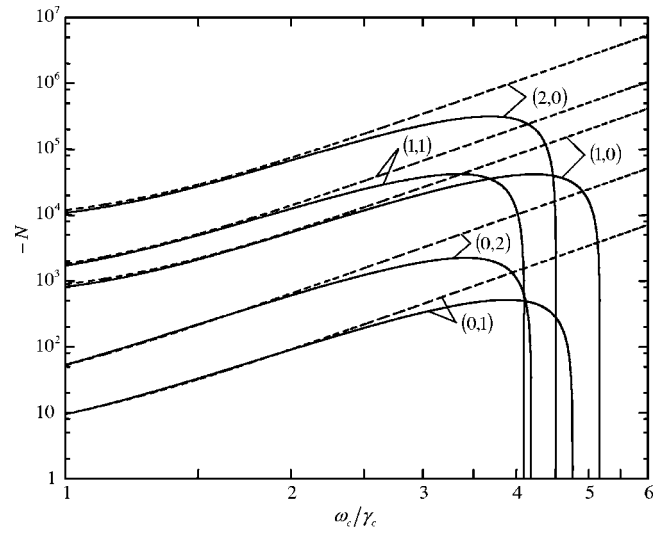


FIG. 5. Negative N regions for FC-71 at two different equilibrium reference states: Graphs of $-N$ for different carrier modes (n_c, m_c) as functions of ω_c/γ_c in double logarithmic scale. Critical values of FC-71: $\bar{\vartheta}_{cr} = 646.2$ K, $\bar{p}_{cr} = 9.423 \times 10^5$ Pa, $\bar{\rho}_{cr} = 619.3$ kg/m³; dashed lines: $\bar{\vartheta}_0 = \bar{\vartheta}_{cr}$, $\bar{p}_0 = 0.911\bar{p}_{cr}$, $\bar{\rho}_0 = 0.475\bar{\rho}_{cr}$, $\Gamma_0 = 0.0$, $\Lambda_0 = -1.162$; solid lines: $\bar{\vartheta}_0 = \bar{\vartheta}_{cr}$, $\bar{p}_0 = 0.887\bar{p}_{cr}$, $\bar{\rho}_0 = 0.441\bar{\rho}_{cr}$, $\Gamma_0 = 0.101$, $\Lambda_0 = -1.131$.

of state and taking the input parameters from Ref. 23. The first reference state was chosen to be somewhere on the (second) transition line mentioned above where $\Gamma_0 = 0$ and $\Lambda_0 < 0$; the other is located on the same isotherm as the first one, however, it corresponds to a lower density. Similar results obtained for the siloxane MD₆M are displayed in Fig. 6. Here, the values for Γ_0 were taken from Fig. 10 in Ref. 25. The solutions presented there are based on the Peng-Robinson equation of state. Furthermore, for each reference state, Λ_0 was computed by exploiting the fact that BZT fluids have large specific heats and, thus, the isentropes and the

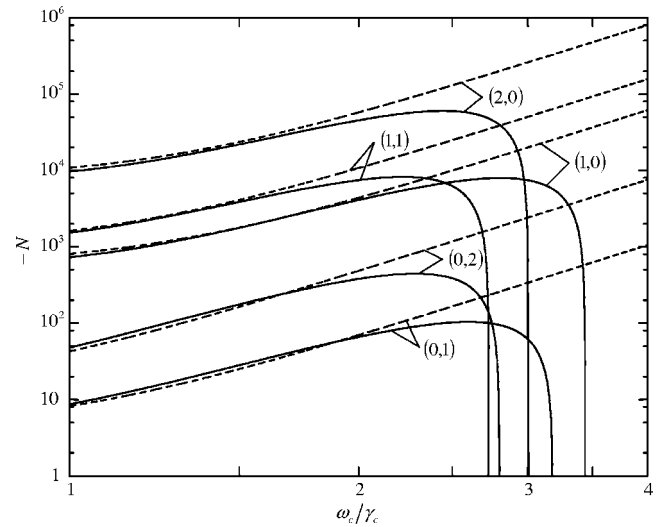


FIG. 6. Negative N regions for MD₆M at two different equilibrium reference states: Graphs of $-N$ for different carrier modes (n_c, m_c) as functions of ω_c/γ_c in double logarithmic scale. Critical values of MD₆M: $\bar{\vartheta}_{cr} = 688.9$ K, $\bar{p}_{cr} = 6.770 \times 10^5$ Pa, $\bar{\rho}_{cr} = 246.1$ kg/m³; dashed lines: $\bar{\vartheta}_0 = \bar{\vartheta}_{cr}$, $\bar{p}_0 = 0.975\bar{p}_{cr}$, $\bar{\rho}_0 = 0.641\bar{\rho}_{cr}$, $\Gamma_0 \approx 0.0$, $\Lambda_0 \approx -0.87$; solid lines: $\bar{\vartheta}_0 = \bar{\vartheta}_{cr}$, $\bar{p}_0 = 0.946\bar{p}_{cr}$, $\bar{\rho}_0 = 0.555\bar{\rho}_{cr}$, $\Gamma_0 \approx 0.15$, $\Lambda_0 \approx -0.95$.

isothermes in the pressure-density state space almost coincide.²² As a consequence, the isentropic derivative of Γ with respect to the density [see Eq. (3)] can roughly be estimated from its isothermal derivative.

C. Solutions

In the (unreal) case of vanishing dissipation, i.e., $\sigma=0$, Eq. (125) can easily be shown to be satisfied by the so-called *continuous wave* (CW) solution

$$\hat{U}_{cw} = a e^{j(w\xi - (w^2/2)Z - \text{sgn}(N)a^2Z + \phi_0)}, \quad (132)$$

with the real parameters a and ϕ_0 being the amplitude and the initial phase, respectively. The arbitrary real constant w characterizes the frequency shift due to dispersion. It is well known that if the nonlinearity parameter N is negative, a solution of this form is *modulationally unstable* to small periodic perturbations $\propto \exp(j\Omega\xi)$, provided that the modulation frequencies Ω are selected in a finite band $-2a < \Omega < 2a$ (see, for example, Ref. 33). In other words, in this case the nonlinearity is obviously no longer able to balance the dispersion and one can expect the generation of new stable solutions. Indeed, by applying the method of the *inverse scattering transform* (IST) such a solution called the *bright* NLS soliton,

$$N < 0: \quad \hat{U}_{bs} = a \text{sech}(a(\xi - wZ - \xi_0)) \times e^{j(w\xi - (w^2/2)Z + (a^2/2)Z + \phi_0)}, \quad (133)$$

can be found,³⁴ where the parameter ξ_0 stands for the initial position of the soliton and a and w now also play the roles as measures for the inverse pulse width and the inverse pulse velocity relative to the inverse group velocity ω_g , respectively. The condition $N < 0$ is thus said to define the type of the *focusing* cubic nonlinearity. As explicated in Sec. IV B, this case is of great importance if the evolution of a wave packet through a BZT fluid is considered. The properties of solitons have been studied extensively in the past; a detailed discussion of their most significant features and the extension to the multiple-soliton solution is presented, e.g., in Ref. 35 (Chap. 1 and pp. 182–183). Other exact solutions to the focusing NLS equation of practical relevance can be found, for instance, in Ref. 36 and the references therein.

In contrast to that, if $N > 0$, the result (132) proves to be *modulationally stable* to spatial periodic perturbations¹⁸ and the problem is said to be governed by the *defocusing* version of the NLS equation (125). Worth mentioning is the fact that in this case solutions in the form of “dark” pulses, i.e., rapid dips in the intensity, created on the CW background are also possible.³⁷ The general form of these so-called *dark* NLS solitons corresponding to the boundary condition $|\hat{U}| \rightarrow a$ at $\xi \rightarrow \pm\infty$ is

$$N > 0: \quad \hat{U}_{ds} = (c - jd \tanh(ad[\xi - (w + ac)Z - \xi_0])) \hat{U}_{cw}, \quad c^2 + d^2 = 1, \quad (134)$$

where the independent parameter d characterizes the soliton intensity. However, as brought out in the literature on pulse propagation in optical fibers (see, e.g., Ref. 38), the generation of dark solitons in an experimental setup is a rather

difficult task since it requires the background wave to be finely tuned to satisfy their demand for a local change in phase and amplitude.

In order to derive the wave number shift for purely monochromatic waves, i.e., single frequency waves without temporal modulation, that is induced by the cubic nonlinearity, the parameter w appearing in the CW solution (132) simply has to be set to 0. Returning to the original coordinate z by using Eq. (124), one then obtains the following expression^{16,18} for the velocity potential Φ_1 from Eq. (114):

$$N > 0: \quad \Phi_1 = A \frac{J_{m_c}(\gamma_c r)}{J_{m_c}(\gamma_c)} e^{j[\omega_c t - (k_c + k_{nl})z - m_c \phi - \phi_0]} + CC, \quad (135)$$

$$k_{nl} = -M^2 A^2 \frac{N}{k_c},$$

where k_{nl} and $A = a\sqrt{k_c}/N$ denote, respectively, the nonlinear wave number shift and the amplitude of the leading order perturbation.

For the type of waves considered in the present study, the dissipative effects cannot be neglected. As a consequence, the acoustic motion in the duct will be governed by the full version of the NLS equation (125), provided that the condition $M \ll \sigma \ll O(1)$ is satisfied. Equivalently, the set of Eqs. (127) and (128) can be used if the damping parameter is so small that $M \ll \sigma \ll 1$ holds. Moreover, please note that once the solutions for \hat{U} and thus Φ_1 are obtained, the expression for Φ_M can readily be derived from Eq. (117).

Finding the damped CW solution for the defocusing case is attempted by substitution of the ansatz $\hat{U} = a(Z) \exp(j(w\xi - w^2Z/2 + \phi(Z)))$ into Eq. (125), whereupon integration leads to

$$N > 0: \quad \hat{U}_{cwd} = a_0 e^{-\sigma DZ + j[w\xi - (w^2/2)Z + (a_0^2/2\sigma D)(e^{-2\sigma DZ} - 1) + \phi_0]}. \quad (136)$$

Worth emphasizing is the fact that this is an asymptotically correct result for all values of σ within the range $M \ll \sigma \ll O(1)$, and it reduces to the unperturbed solution when formally the limit $\sigma \rightarrow 0$ is taken. The expression for the velocity potential of a purely monochromatic wave thus reads

$$N > 0: \quad \Phi_1 = A_0 \frac{J_{m_c}(\gamma_c r)}{J_{m_c}(\gamma_c)} e^{-\hat{\Gamma}_c z + j(\omega_c t - \int_0^z k_{nl} d^2 z - m_c \phi - \phi_0)} + CC, \quad (137)$$

$$\hat{\Gamma}_c = jk_c + \alpha jk_c \hat{F}_c, \quad k_{nl} = -M^2 A_0^2 \frac{N}{k_c} e^{-2\alpha k_c \Re(\hat{F}_c)z}.$$

Here, k_{nl} and $A_0 = a_0\sqrt{k_c}/N$ denote, respectively, the *local* nonlinear wave number shift under the action of damping and the amplitude of Φ_1 at $z=0$, and \hat{F}_c is the quantity defined in Eq. (58), with $(n, m) = (n_c, m_c)$. According to what was stated before, the perturbation parameter $\alpha = \sigma M^2 \sim \text{Re}^{-1/2}$ has to be chosen such that $M^3 \ll \alpha \ll O(M^2)$. As expected, in the limit as $z \rightarrow \infty$ the nonlinear effects are damped away and the result from Eq. (137) ap-

proaches the solution for linear sound propagation truncated after the first correction term [see Eq. (81)].

The subject of bright and dark soliton dynamics in weakly dissipative or otherwise perturbed systems that can be described by an extended NLS equation has been attracting considerable interest over the last 30 years (see, e.g., the review papers by Kivshar *et al.*^{38,39}). In particular, recent progress in communication systems technology has presented many research problems in connection with pulse propagation in optical fibers where, similar to the problem studied here, the soliton represents the envelope of a wave packet. The analytical description of solitons under the action of (small) perturbations can be approached by different methods. Analyzing the prototype case of linear damping (125), Nicholson and Goldman⁴⁰ derived the general expressions for the evolution of the bright soliton parameters by exploiting the modified conservation laws of the perturbed NLS equation. Theories based on the IST for $N < 0$ were introduced by Kaup⁴¹ and independently also by Karpman and Maslov,⁴² whereas Keener and McLaughlin⁴³ and Ablowitz and Segur³⁵ (pp. 270–271) obtained equivalent results by solving the multiple scales secularity conditions of the higher order problem [here given by Eq. (128)] that derive from the Fredholm alternative or related concepts. These so-called direct perturbation techniques and the IST methods have the advantage that, in addition to the evolution of the leading order soliton solution, the effects arising in the higher orders in σ can be taken into account as well.

Here, a more intuitive approach strongly related to the technique applied in Sec. III shall be used in order to reveal the most significant features of a bright soliton propagating in the system given by Eqs. (127) and (128): The solvability conditions are derived by insisting on \hat{U}_σ not having worse behavior as Z and ξ tend to infinity than \hat{U}_1 . To this end, solutions of the form

$$\begin{aligned}\hat{U}_1 &= a(Z_\sigma) \hat{f}_1(Z, Z_\sigma, \xi) e^{j(w\xi - (w^2/2)Z)}, \\ \hat{U}_\sigma &= a(Z_\sigma) \hat{f}_\sigma(Z, Z_\sigma, \xi) e^{j(w\xi - (w^2/2)Z)}\end{aligned}\quad (138)$$

are sought. Moreover, the following coordinate transformation

$$\begin{aligned}s(Z) &= \int_0^{\sigma Z} \frac{a^2(Z_\sigma)}{\sigma} dZ_\sigma + s_0, \\ \eta(Z, Z_\sigma, \xi) &= a(Z_\sigma)(\xi - wZ - \xi_0), \quad \bar{Z}_\sigma(Z_\sigma) = Z_\sigma\end{aligned}\quad (139)$$

is applied, which requires the real-valued function $a(Z_\sigma)$ to be determined such that s and η are order $O(1)$ quantities within the spatial and temporal ranges defined by $Z \sim z_2 = O(1)$ and $\xi \sim (\omega_g t_1 - z_1) = O(1)$ and, thus, can take over the roles of the “shorter” scales formerly played by Z and ξ . It should be mentioned that a very similar procedure was used in Crighton *et al.*²⁸ (pp. 214–219) in order to analyze the effect of gradual sound speed variation on acoustic waves. Having substituted Eq. (138) into Eqs. (127) and (128), the transformation (139) leads to

$$j \frac{\partial \hat{f}_1}{\partial s} + \frac{1}{2} \frac{\partial^2 \hat{f}_1}{\partial \eta^2} - \text{sgn}(N) \hat{f}_1 \hat{f}_1^* = 0, \quad (140)$$

$$\begin{aligned}j \frac{\partial \hat{f}_\sigma}{\partial s} + \frac{1}{2} \frac{\partial^2 \hat{f}_\sigma}{\partial \eta^2} - \text{sgn}(N) \hat{f}_1 (2 \hat{f}_1^* \hat{f}_\sigma + \hat{f}_1 \hat{f}_\sigma^*) \\ = -j \frac{1}{a^2} \left(\frac{\partial \hat{f}_1}{\partial \bar{Z}_\sigma} + D \hat{f}_1 \right) - j \frac{1}{a^3} \frac{da}{d\bar{Z}_\sigma} \left(\eta \frac{\partial \hat{f}_1}{\partial \eta} + \hat{f}_1 \right).\end{aligned}\quad (141)$$

Obviously, the solutions to Eq. (140) are subject to the same stability conditions that hold for the solutions of the undamped case discussed before. Hence, the function \hat{f}_1 is taken to be

$$N < 0: \quad \hat{f}_1 = \text{sech}(\eta) e^{js/2}. \quad (142)$$

Since, furthermore, $\partial \hat{f}_1 / \partial s$ also constitutes a solution to the homogeneous part of Eq. (141), a particular solution can be expected to be found by making the ansatz

$$\hat{f}_\sigma = h(s, \bar{Z}_\sigma, \eta) \frac{\partial \hat{f}_1}{\partial s}, \quad (143)$$

where h is a real-valued function. Separating real and imaginary parts of the terms that are generated by substituting Eq. (143) into Eq. (141) and solving the resulting equations yields $\partial h / \partial s = 0$ and

$$\begin{aligned}h = -\frac{\eta^2}{a^3} \frac{da}{d\bar{Z}_\sigma} + \frac{\eta}{a^3} \left(\frac{da}{d\bar{Z}_\sigma} + 2Da + 2h_1(\bar{Z}_\sigma) \right) \\ - \frac{e^{-2\eta}}{2a^3} \left(\frac{da}{d\bar{Z}_\sigma} + 2Da + h_1(\bar{Z}_\sigma) \right) + \frac{e^{2\eta}}{2a^3} h_1(\bar{Z}_\sigma) + h_2(\bar{Z}_\sigma).\end{aligned}\quad (144)$$

The terms in Eq. (145) proportional to $\exp(2\eta)$ and $\exp(-2\eta)$ obviously have to be annihilated since they would involve secular terms in \hat{f}_σ . This implies that $h_1 = 0$ and

$$a = a_0 e^{-2D\bar{Z}_\sigma}. \quad (146)$$

Consequently, the leading order solution for the damped bright soliton and the particular solution for its first-order correction read

$N < 0, \quad M \ll \sigma \ll 1:$

$$\begin{aligned}\hat{U}_{1bsd} &= a_0 \text{sech}(a_0 e^{-2\sigma DZ} (\xi - wZ - \xi_0)) \\ &\times e^{-2\sigma DZ + j[w\xi - (w^2/2)Z - (a_0^2/8\sigma D)(e^{-4\sigma DZ} - 1) + \phi_0]}, \\ \hat{U}_{\sigma bsd} &= -jD(\xi - wZ - \xi_0)^2 \hat{U}_{1bsd}.\end{aligned}\quad (147)$$

These solutions are in complete conformity with the results from Refs. 35 and 40–43, and, utilizing Eqs. (114) and (124), they can easily be recast into the expressions for the velocity potentials $\Phi_1 \propto \hat{U}_1$ and $\Phi_\sigma \propto \hat{U}_\sigma$; the next higher order correction Φ_M is then obtained from Eq. (117). It should be emphasized that this combined solution is valid over spatial and temporal ranges of the order $O(\sigma^{-1} M^{-1})$ only.

As far as the propagation of dark solitons for the damped defocusing case is concerned, the study by Lisak *et al.*⁴⁴ based upon the modified conservation laws and the more refined analysis by Kivshar *et al.*³⁸ using the IST per-

turbation theory should be mentioned. To the author's knowledge, the most complete investigation is presented in the 1998 paper by Chen *et al.*,⁴⁵ who, applying a direct perturbation technique to the system (127) and (128), arrived at the relationships $a(\bar{Z}_\sigma) = a_0 \exp(-D\bar{Z}_\sigma)$ and, additionally, $\xi_0(\bar{Z}_\sigma) = \xi_{00} + D\bar{Z}_\sigma/(2a_0d)$. Thus, the expression for the damped dark soliton assumes the form

$N > 0, \quad M \ll \sigma \ll 1:$

$$\hat{U}_{1dsd} = (c - jd \tanh(\beta)) \hat{U}_{c wd}, \quad c^2 + d^2 = 1, \quad (148)$$

$$\beta = a_0 e^{-\sigma D Z} d \left[\xi - (w + a_0 e^{-\sigma D Z} c) Z - \xi_{00} \right] - \frac{\sigma D Z}{2} e^{-\sigma D Z},$$

in which $U_{c wd}$ is the solution for the damped background wave (136). The (very complex) result for the first-order correction $U_{\sigma dsd}$ can also be found in Ref. 45.

Despite being valid for a wide range of wave numbers and frequencies, the solvability condition for the third-order problem in the form (123) obviously breaks down in the limit as $k_c \rightarrow 0$. That is to say, the solutions above cease to be valid for driving frequencies near the linear cutoff frequency of the carrier mode. However, this difficulty can be overcome by using instead of Eq. (123) the more general conditions (116) and (118). With k_c set to zero, they read $\partial \hat{A}_1 / \partial t_1 = 0$ and

$$2j\gamma_c \frac{\partial \hat{A}_1}{\partial t_2} - \frac{\partial^2 \hat{A}_1}{\partial z_1^2} = -2N \hat{A}_1^2 \hat{A}_1^* + 2\sigma \hat{F}_c \hat{A}_1, \quad (149)$$

respectively. Here, \hat{F}_c is the parameter defined by Eq. (83), with $\omega_{\bar{m}} = \omega_c = \gamma_c$ and $(\bar{n}, \bar{m}) = (n_c, m_c)$. Introduction of the scaled quantities

$$Z = M^2 t, \quad \xi = M \sqrt{\gamma_c} z, \quad D = \frac{\Re(\hat{F}_c)}{\gamma_c} > 0, \quad (150)$$

$$\hat{U} = \sqrt{\frac{|N|}{\gamma_c}} \hat{A}_1^* e^{-j\sigma D Z}$$

then again yields the damped NLS equation (125).

Please note that upon neglecting the nonlinear terms in Eqs. (149) and recalling that $\sigma = \alpha/M^2$, the solution for Φ_1

$$\Phi_1 = A \frac{J_{m_c}(\gamma_c r)}{J_{m_c}(\gamma_c)} e^{-\hat{\Gamma}_c z + j(\Omega_c t - m_c \phi - \phi_0)} + \text{CC}, \quad (151)$$

$$\hat{\Gamma}_c = \alpha^{1/2} j(2\hat{F}_c + 2\gamma_c \omega_\alpha)^{1/2}, \quad \Omega = \gamma_c + \alpha \omega_\alpha,$$

where $\omega_\alpha = O(1)$ is an arbitrary detuning frequency, agrees perfectly with the linear solution from Eq. (82) truncated after the leading order term. On the other hand, if the viscous effects are neglected, the CW solution (132) for $N > 0$ with $w = 0$ gives Nayfeh's solution¹⁸ for the nonlinear shift ω_{nl} of the cutoff frequency of a monochromatic wave:

$$N > 0, \quad \sigma = 0: \quad \Phi_1 = A \frac{J_{m_c}(\gamma_c r)}{J_{m_c}(\gamma_c)} e^{j[(\gamma_c + \omega_{nl})t - m_c \phi - \phi_0]} + \text{CC}, \quad (152)$$

$$\omega_{nl} = M^2 A^2 \frac{N}{\gamma_c}.$$

However, inspection of Eq. (136) immediately shows that in the presence of boundary layer effects, the concept of using the (damped) CW solution in order to find analytical expressions for Φ_1 as $k_c \rightarrow 0$ and the corresponding boundary condition at $z = 0$ fails since any such solution, propagating ($w \neq 0$) or not propagating ($w = 0$), would involve the term $\exp(-\alpha \Re(\hat{F}_c) t / \gamma_c)$ and, thus, would become unbounded as $t \rightarrow -\infty$. Notwithstanding, if $N > 0$, the damped NLS equation (125) can be solved numerically for any bounded boundary condition. In contrast to Eq. (136), if $N < 0$, $M \ll \sigma \ll 1$, and $k_c \rightarrow 0$, the damped soliton solution (147) in combination with Eqs. (114) and (150) *does* provide a valid solution for Φ_1 that is bounded for all values of t and x .

V. CONCLUSIONS

The work presented here analyzes the linear and the nonlinear propagation of sound waves in circular, hard-walled tubes containing a real gas. The investigations primarily proceed from the assumption that the acoustic boundary layer is thin compared to the diameter of the duct and that the characteristic wavelength and the diameter are of the same order of magnitude. In the first part, dealing with the linear problem, the method of matched asymptotic expansions in combination with the method of multiple scales has been applied in order to derive the first- and second-order terms of the sound pressure and the expansion of the propagation parameters up to third order. These solutions generalize the results given in the literature for real gases and for the cases where the tube length becomes so large that the effects resulting from bulk viscosity and heat conduction in the core region have to be taken into account as well.

The second part deals with the nonlinear evolution of a wave packet centered at the frequency and the wave number of a higher order carrier mode. Using the same perturbation methods as before, the damped nonlinear Schrödinger equation (125) has been derived, which proves to be applicable even if the driving frequency of the carrier mode is close to its cutoff frequency. Furthermore, special emphasis has been placed on a detailed discussion of the sign of the nonlinearity parameter N : In the case of a perfect gas the nonlinearity parameter is a strictly positive quantity and, thus, the evolution of the wave packet is governed by the defocusing version of the Schrödinger equation. In contrast to that, for special combinations of the parameters characterizing the equilibrium reference state of a real gas, also the focusing version can be obtained, enabling the formation of stable soliton solutions. Particularly, the so-called BZT fluids have turned out to be possible candidates for applications where the nonlinearity parameter becomes negative.

The asymptotic analysis which has led to Eq. (125), obviously, assumes that $N = O(1)$ and will break down in the

vicinity of a transition points where the nonlinear parameter vanishes; that such transition points might occur can be seen from Figs. 5 and 6. For values of ω_c away from these points, the wave evolution will be described by Eq. (125) with either $N > 0$ or $N < 0$. However, within a $N = O(M)$ neighborhood of these points, small terms neglected in constructing Eq. (125) become of the same order as the already included nonlinear correction term and a separate analysis is then required. In a different physical context, asymptotic expansions to analyze this case have been developed by Grimshaw *et al.*³⁶ and Parkes.⁴⁶ Neglecting all dissipative effects, they arrived at a modified nonlinear Schrödinger equation featuring additional cubic and quintic nonlinearity terms. It would be interesting to know whether a similar evolution equation that also incorporates the boundary layer effects can be derived for the problem studied here.

ACKNOWLEDGMENT

The author would like to thank Marcus Wrabel at the Institute of Fluid Mechanics and Heat Transfer, Vienna University of Technology, for providing the numerical code to evaluate the Martin-Hou equation of state for the BZT fluid FC-71.

APPENDIX: PROPAGATION PARAMETERS IN DIMENSIONAL FORM

Please note that here the accents “ \sim ” introduced in Sec. I to indicate dimensional quantities have been omitted. Equation (80):

$$\hat{\Gamma}_{nm} = \sqrt{-\hat{k}_{nm}^2 \left(1 + 2 \frac{\hat{F}_{nm}}{\text{St}(\omega)} + \frac{2\hat{H}_{nm} + \hat{F}_{nm}^2}{\text{St}(\omega)^2} \right) + O(\text{St}^{-3})}, \quad (\text{A1})$$

where

$$\hat{k}_{mn} = \begin{cases} 1/R \sqrt{\text{He}(\omega)^2 - \gamma_{nm}^2}, & \text{He}(\omega) > \gamma_{nm}, \\ -j/R \sqrt{\gamma_{nm}^2 - \text{He}(\omega)^2} & \text{He}(\omega) < \gamma_{nm}, \end{cases}$$

$$\hat{F}_{nm} = \frac{1-j}{-\sqrt{2}(1-m^2/\gamma_{nm}^2)} \times \left(1 + \frac{m^2}{R^2 \hat{k}_{nm}^2} + \frac{\text{He}(\omega)^2}{R^2 \hat{k}_{nm}^2} \frac{\gamma_0 - 1}{\sqrt{\text{Pr}}} \right),$$

$$\hat{H}_{00} = -j \left[1 + \frac{\gamma_0 - 1}{\sqrt{\text{Pr}}} \left(1 - \frac{\gamma_0}{2\sqrt{\text{Pr}}} \right) \right] - \frac{j\text{He}(\omega)^2}{2} \left[\frac{11}{6} + \frac{\eta_0}{\mu_0} + \frac{\gamma_0 - 1}{2\sqrt{\text{Pr}}} \left(2 + \frac{1 + \gamma_0}{\sqrt{\text{Pr}}} \right) \right],$$

$$\hat{H}_{nm} = \hat{G}_{nm} - \frac{\hat{F}_{nm}^2}{2} + \frac{R^2 \hat{k}_{nm}^2 \hat{F}_{nm}^2}{\gamma_{nm}^2} \frac{m^2}{\gamma_{nm}^2 - m^2} - \frac{j\text{He}(\omega)^4}{2\hat{k}_{nm}^2 R^2} \left(\frac{4}{3} + \frac{\eta_0}{\mu_0} + \frac{\gamma_0 - 1}{\text{Pr}} \right), \quad (n, m) \neq (0, 0),$$

$$\hat{G}_{nm} = -\frac{j}{2(1-m^2/\gamma_{nm}^2)^2} \left[3 + \frac{m^2}{\gamma_{nm}^2} + \frac{m^2}{R^2 \hat{k}_{nm}^2} \left(5 - \frac{m^2}{\gamma_{nm}^2} \right) - \frac{\text{He}(\omega)^2}{R^2 \hat{k}_{nm}^2} (\gamma_0 - 1) \left(\left(1 - \frac{m^2}{\gamma_{nm}^2} \right) \frac{1}{\text{Pr}} - \frac{4}{\sqrt{\text{Pr}}} \right) \right],$$

$$(n, m) \neq (0, 0),$$

$$\text{St}(\omega) = \sqrt{\frac{\omega \rho_0 R^2}{\mu_0}} \gg 1, \quad \text{He}(\omega) = \frac{\omega R}{c_0} = O(1),$$

$$\text{Pr} = \frac{\mu_0 C_p}{\kappa_0} = O(1). \quad (\text{A2})$$

Equation (94):

$$\hat{\Gamma}_{nm} = \frac{j\text{He}(\omega)}{R} \left(1 + \frac{\hat{L}_{nm}}{\text{St}(\omega)} + O(\text{St}^{-2}) \right), \quad (\text{A3})$$

where

$$\hat{L}_{nm} = -\frac{\hat{\gamma}_{nm}^2 \text{St}(\omega)}{2\text{He}(\omega)^2} - \frac{j\text{He}(\omega)^2}{2\text{St}(\omega)} \left(\frac{4}{3} + \frac{\eta_0}{\mu_0} + \frac{\gamma_0 - 1}{\text{Pr}} \right),$$

$$\hat{J}_{m+1}(\hat{\gamma}_{nm}) \hat{\gamma}_{nm} - m \hat{J}_m(\hat{\gamma}_{nm}) + \frac{\text{He}(\omega)^2 \hat{F}}{\text{St}(\omega)} \hat{J}_m(\hat{\gamma}_{nm}) = 0,$$

$$\hat{F} = \frac{1-j}{\sqrt{2}} \left(1 + \frac{\gamma_0 - 1}{\sqrt{\text{Pr}}} \right),$$

$$\text{St}(\omega) = \sqrt{\frac{\omega \rho_0 R^2}{\mu_0}} \gg 1, \quad \text{He}(\omega) = \frac{\omega R}{c_0} = O(\text{St}^{1/2}),$$

$$\text{Pr} = \frac{\mu_0 C_p}{\kappa_0} = O(1). \quad (\text{A4})$$

¹G. Kirchhoff, “Über den Einfluß der Wärmeleitung in einem Gase auf die Schallbewegung (On the effect of heat conduction in a gas on the propagation of sound),” (in German). *Ann. Phys.* **134**, 177–193 (1868).

²L. Rayleigh, *Theory of Sound II* (Macmillan, London, 1896), pp. 319–328.

³C. Zwikker and C. Kosten, *Sound Absorbing Materials* (Elsevier, Amsterdam, 1949), pp. 25–40.

⁴H. Tijdeman, “On the propagation of sound waves in cylindrical tubes,” *J. Sound Vib.* **39**(1), 1–33 (1975).

⁵D. H. Keefe, “Acoustical wave propagation in cylindrical ducts: Transmission line parameter approximations for isothermal and nonisothermal boundary conditions,” *J. Acoust. Soc. Am.* **75**, 58–62 (1984).

⁶S. Scheichl, “On the calculation of the transmission line parameters for long tubes using the method of multiple scales,” *J. Acoust. Soc. Am.* **115**, 534–555 (2004).

⁷R. E. Beatty, Jr., “Boundary layer attenuation of higher order modes in rectangular and circular tubes,” *J. Acoust. Soc. Am.* **22**, 850–854 (1950).

⁸A. M. Bruneau, M. Bruneau, Ph. Herzog, and J. Kergomard, “Boundary layer attenuation of higher order modes in waveguides,” *J. Sound Vib.* **119**, 15–27 (1987).

⁹H. Hudde, “The propagation constant in lossy circular tubes near the cutoff frequencies of higher-order modes,” *J. Acoust. Soc. Am.* **83**, 1311–1318 (1988).

¹⁰W. Chester, “Resonant oscillations in closed tubes,” *J. Fluid Mech.* **18**, 44–64 (1964).

¹¹D. T. Blackstock, “Nonlinear acoustics (theoretical),” in *American Institute of Physics Handbook*, edited by D. Gray (McGraw–Hill, New York, 1972), Chap. 3, pp. 183–205.

¹²V. Gusev, H. Bailliet, P. Lotton, and M. Bruneau, “Asymptotic theory of

- nonlinear acoustic waves in a thermoacoustic prime-mover," *Acustica* **86**, 25–38 (2000).
- ¹³ L. Menguy and J. Gilbert, "Weakly nonlinear gas oscillations in air-filled tubes; solutions and experiments," *Acustica* **86**, 798–810 (2000).
- ¹⁴ S. N. Makarov and E. V. Vatrushina, "Effect of the acoustic boundary layer on a nonlinear quasiplane wave in a rigid-walled tube," *J. Acoust. Soc. Am.* **94**, 1076–1083 (1993).
- ¹⁵ S. Makarov and M. Ochmann, "Nonlinear and thermoviscous phenomena in acoustics, part II," *Acustica* **83**, 197–222 (1997).
- ¹⁶ J. B. Keller and M. H. Millman, "Finite-amplitude sound-wave propagation in a waveguide," *J. Acoust. Soc. Am.* **49**, 329–333 (1971).
- ¹⁷ M. A. Foda, "Analysis of nonlinear propagation an interactions of higher order modes in a circular waveguide," *Acustica* **84**, 66–77 (1998).
- ¹⁸ A. H. Nayfeh, "Nonlinear propagation of a wave packet in a hard-walled circular duct," *J. Acoust. Soc. Am.* **57**, 803–809 (1975).
- ¹⁹ H. A. Bethe, "The theory of shock waves for an arbitrary equation of state," *Office Sci. Res. Dev. Rep.* **545**, (1942).
- ²⁰ Ya. B. Zel'dovich, "On the possibility of rarefaction shock waves," *Zh. Eksp. Teor. Fiz.* **4**, 363–364 (1946).
- ²¹ P. A. Thompson, "A fundamental derivative in gasdynamics," *Phys. Fluids* **14**, 1843–1849 (1971).
- ²² P. A. Thompson and K. C. Lambrakis, "Negative shock waves," *J. Fluid Mech.* **60**, 187–207 (1973).
- ²³ M. S. Cramer, "Negative nonlinearity in selected fluorocarbons," *Phys. Fluids A* **1**, 1894–1897 (1989).
- ²⁴ J. F. Monaco, M. S. Cramer, and L. T. Watson, "Supersonic flows of dense gases in cascade configurations," *J. Fluid Mech.* **330**, 31–59 (1997).
- ²⁵ P. Colonna and P. Silva, "Dense gas thermodynamic properties of single and multicomponent fluids for fluid dynamics simulations," *J. Fluids Eng.* **125**, 414–427 (2003).
- ²⁶ A. Kluwick, "Small-amplitude finite-rate waves in fluids having both positive and negative nonlinearity," in *Nonlinear Waves in Real Fluids*, edited by A. Kluwick (Springer, Vienna, 1991), pp. 1–43.
- ²⁷ R. B. Bird and M. D. Graham, "General equations of Newtonian fluid dynamics," in *The Handbook of Fluid Dynamics*, edited by R. W. Johnson (CRC, Boca Raton, and Springer, Heidelberg, 1998), Chap. 3, p. 17.
- ²⁸ D. G. Crighton, A. P. Dowling, J. E. Ffowcs Williams, M. Heckl, and F. G. Leppington, *Modern Methods in Analytical Acoustics* (Springer, Berlin, 1996).
- ²⁹ A. H. Nayfeh and D. P. Telionis, "Acoustic propagation in ducts with varying cross sections," *J. Acoust. Soc. Am.* **54**, 1654–1661 (1973).
- ³⁰ J. Kergomard, M. Bruneau, A. M. Bruneau, and Ph. Herzog, "On the propagation constant of higher order modes in a cylindrical waveguide," *J. Sound Vib.* **126**, 178–181 (1988).
- ³¹ M. Bruneau, Ph. Herzog, J. Kergomard, and J. D. Polack, "General formulation of the dispersion equation in bounded visco-thermal fluid, and application to some simple geometries," *Wave Motion* **11**, 441–451 (1989).
- ³² T. Taniuti, "Reductive perturbation method and far fields of wave equations," *Prog. Theor. Phys. Suppl.* **55**, 1–35 (1974).
- ³³ A. Hasegawa, *Optical Solitons in Fibers* (Springer, Berlin, 1989), pp. 56–62.
- ³⁴ V. E. Zakharov and A. B. Shabat, "Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media," *Sov. Phys. JETP* **34**, 62–69 (1972).
- ³⁵ M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform* (SIAM, Philadelphia, 1981).
- ³⁶ R. Grimshaw, D. Pelinovsky, E. Pelinovsky, and T. Talipova, "Wave group dynamics in weakly nonlinear long-wave models," *Physica D* **159**, 35–57 (2001).
- ³⁷ V. E. Zakharov and A. B. Shabat, "Interactions between solitons in a stable medium," *Sov. Phys. JETP* **37**, 823–828 (1973).
- ³⁸ Y. S. Kivshar and B. Luther-Davies, "Dark optical solitons: physics and applications," *Phys. Rep.* **298**, 81–197 (1998).
- ³⁹ Y. S. Kivshar and B. A. Malomed, "Dynamics of solitons in nearly integrable systems," *Rev. Mod. Phys.* **61**, 763–915 (1989).
- ⁴⁰ D. R. Nicholson and M. D. Goldman, "Damped nonlinear Schrödinger equation," *Phys. Fluids* **19**, 1621–1625 (1976).
- ⁴¹ D. J. Kaup, "A perturbation expansion for the Zakharov-Shabat inverse scattering transform," *SIAM (Soc. Ind. Appl. Math.) J. Appl. Math.* **31**, 121–133 (1976).
- ⁴² V. I. Karpman and E. M. Maslov, "Perturbation theory for solitons," *Sov. Phys. JETP* **46**, 281–291 (1977).
- ⁴³ J. P. Keener and D. W. McLaughlin, "Solitons under perturbation," *Phys. Rev. A* **16**, 777–790 (1977).
- ⁴⁴ M. Lisak, D. Anderson, and B. A. Malomed, "Dissipative damping of dark solitons in optical fibers," *Opt. Lett.* **16**, 1936–1937 (1991).
- ⁴⁵ X.-J. Chen, Z.-D. Chen, and N.-N. Huang, "A direct perturbation theory for dark solitons based on a complete set of squared Jost solutions," *J. Phys. A* **31**, 6929–6947 (1998).
- ⁴⁶ E. J. Parkes, "The modulation of weakly non-linear dispersive waves near the marginal state of instability," *J. Phys. A* **20**, 2025–2036 (1987).

Simultaneous measurement of acoustic and streaming velocities in a standing wave using laser Doppler anemometry^{a)}

Michael W. Thompson^{b)} and Anthony A. Atchley^{c)}

Graduate Program in Acoustics, The Pennsylvania State University, University Park, Pennsylvania 16802

(Received 1 August 2004; revised 27 November 2004; accepted 29 December 2004)

Laser Doppler anemometry (LDA) with burst spectrum analysis (BSA) is used to study the acoustic streaming generated in a cylindrical standing-wave resonator filled with air. The air column is driven sinusoidally at a frequency of approximately 310 Hz and the resultant acoustic-velocity amplitudes are less than 1.3 m/s at the velocity antinodes. The axial component of fluid velocity is measured along the resonator axis, across the diameter, and as a function of acoustic amplitude. The velocity signals are postprocessed using the Fourier averaging method [Sonnenberger *et al.*, Exp. Fluids **28**, 217–224 (2000)]. Equations are derived for determining the uncertainties in the resultant Fourier coefficients. The time-averaged velocity-signal components are seen to be contaminated by significant errors due to the LDA/BSA system. In order to avoid these errors, the Lagrangian streaming velocities are determined using the time-harmonic signal components and the arrival times of the velocity samples. The observed Lagrangian streaming velocities are consistent with Rott's theory [N. Rott, Z. Angew. Math. Phys. **25**, 417–421 (1974)], indicating that the dependence of viscosity on temperature is important. The onset of streaming is observed to occur within approximately 5 s after switching on the acoustic field. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1861233]

PACS numbers: 43.25.Nm, 43.25.Zx [MFH]

Pages: 1828–1838

I. INTRODUCTION

Rayleigh streaming¹ is a form of time-averaged fluid flow that is generated when an acoustic wave in a fluid interacts viscously with a solid surface. The streaming associated with a one-dimensional monofrequency acoustic standing wave in a resonator has previously been investigated by Sharpe *et al.*,² Arroyo and Greated,³ Hann and Greated,⁴ and Rockliff *et al.*⁵ using the particle image velocimetry (PIV) technique. The laser Doppler anemometry (LDA) technique has also been applied to the study of streaming by Thompson and Atchley.^{6,7} Although these studies establish the basic validity of the streaming theory, none of them represents a thorough investigation of the phenomenon.

The present study builds on the previous work of Thompson and Atchley by developing a high-precision method for the simultaneous measurement of acoustic and streaming velocities based on the LDA/BSA technique. This method is used to measure the acoustic streaming generated by a relatively low-amplitude acoustic standing wave in a long cylindrical resonator. The results are compared to the theories of Rott⁸ and Qi.⁹

Section II contains a discussion of the relevant stream-

ing theories, and Sec. III contains a discussion of the previous measurements of the phenomenon. Section IV provides an overview of the LDA technique and the burst spectrum analysis (BSA) method¹⁰ for determining fluid velocity, and Sec. V provides an overview of the Fourier averaging (FA) method¹¹ for analyzing the LDA/BSA data. A description of the experimental apparatus and procedure is given in Sec. VI, an extension of the FA method is presented in Sec. VII, and a method for determining the Lagrangian streaming velocity is presented in Sec. VIII. Sections IX and X present a comparison of the measurements of streaming velocity to the theory.

II. THEORY

The theoretical result for the streaming-velocity field generated by a monofrequency acoustic standing wave in a tube of circular cross section was given by Schuster and Matz.¹² They assumed: (1) that the resonator inner radius R is small relative to the acoustic wavelength λ ; (2) that R is large relative to the thickness of the viscous acoustic boundary layer, $\delta_v = \sqrt{2\nu/\omega}$, where ν is the kinematic viscosity of the fluid and ω is the angular acoustic frequency; (3) that thermal effects are negligible; and (4) that terms higher than second order in the acoustic Mach number $M = U_1/c$ are negligible, where U_1 is the amplitude of the acoustic velocity at its antinode and c is the sound speed in the fluid.

Consider a long cylindrical tube filled with air driven sinusoidally at one end and terminated by a rigid end cap at the other. The acoustic- and streaming-velocity fields in this system can be described in two spatial dimensions: (1) an axial coordinate that runs parallel to the tube centerline and (2) a transverse coordinate that runs perpendicular to the cen-

^{a)}Portions of this work were presented in: M. W. Thompson and A. A. Atchley, "Measurement of acoustic and streaming velocities in fluids using laser Doppler anemometry," 146th Meeting of the Acoustical Society of America, Austin, TX, November 2003; and M. W. Thompson, "Measurement of acoustic streaming in a standing wave using laser Doppler anemometry" (Ph.D. dissertation, The Pennsylvania State University, University Park, PA, 2004).

^{b)}Current affiliation: Applied Physics Laboratory, The Johns Hopkins University, Laurel, Maryland 20723.

Electronic mail: michael.wayne.thompson@jhuapl.edu

^{c)}Electronic mail: atchley@psu.edu

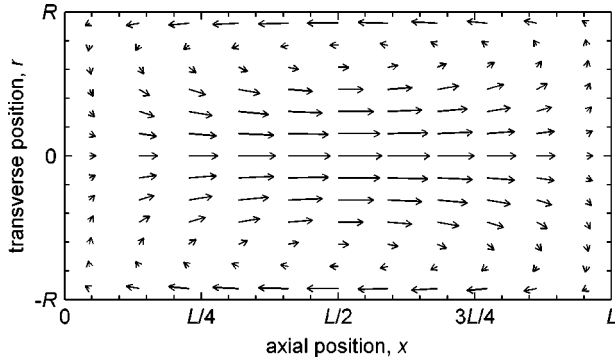


FIG. 1. Profile view of a streaming cell in the outer streaming-velocity field, calculated from Eqs. (2) and (3). The flow along the resonator walls ($r = \pm R$) is toward the acoustic-velocity node ($x=0$), and the return flow along the centerline ($r=0$) is toward the acoustic-velocity antinode ($x=L$). Neighboring cells, to the left and to the right of the cell shown here, circulate in the opposite direction. The variation of the axial component of streaming velocity with respect to x is sinusoidal, with the maximum value occurring at $x=L/2$. The variation of the axial component of streaming velocity with respect to r is parabolic, with the average flow over the cross section (x held constant) equal to zero. As r approaches $\pm R$ within the boundary layer (not shown), the velocity goes to zero.

terline. The velocity fields are assumed to be symmetric about the centerline.

The axial component of the acoustic-velocity field well outside the boundary layer is

$$u_1 = \Re(\tilde{U}_1 e^{j\omega t}) \sin\left(\frac{2\pi x}{\lambda}\right), \quad (1)$$

where \tilde{U}_1 is the complex amplitude of the acoustic velocity at its antinode, t is time, x is the axial coordinate, $j = \sqrt{-1}$, and \Re denotes taking the real part of the complex argument. The transverse component of the acoustic-velocity field well outside the boundary layer is negligible.

The axial component $\langle u_2 \rangle$ and transverse component $\langle v_2 \rangle$ of the streaming-velocity field well outside the boundary layer (the outer streaming-velocity field) are^{12,13}

$$\langle u_2 \rangle = \frac{3}{8} \frac{U_1^2}{c} \left(1 - \frac{2r^2}{R^2}\right) \sin\left(\frac{\pi x}{L}\right), \quad (2)$$

and

$$\langle v_2 \rangle = -\frac{3}{8} \frac{U_1^2}{c} \frac{2\pi r}{\lambda} \left(1 - \frac{r^2}{R^2}\right) \cos\left(\frac{\pi x}{L}\right), \quad (3)$$

where U_1 is the magnitude of \tilde{U}_1 , r is the transverse coordinate ($-R \leq r \leq R$), $L = \lambda/4$, and angled brackets $\langle \dots \rangle$ are used to denote time-averaged quantities. This outer streaming-velocity field consists of a series of toroidal streaming cells of length L , an example of which is shown in Fig. 1. (Within the boundary layer near the resonator walls, there exists an inner streaming-velocity field with vortices circulating in the direction opposite the outer streaming vortices shown here.)

Rott⁸ extended the theory by addressing the effects of heat conduction, including: (1) the acoustic temperature fluctuations; (2) the thermal acoustic boundary layer, with thickness $\delta_\kappa = \sqrt{2\kappa/\omega}$, where κ is the thermal diffusivity of the fluid; (3) the variation of mean temperature T_0 with respect

to the axial coordinate x ; and (4) the dependence of viscosity μ and thermal conductivity on temperature T , assuming a form for the viscosity of $\mu \propto T^\beta$, where β is a constant that depends on the properties of the fluid. [The assumptions of Schuster and Matz¹² numbered (1), (2), and (4) in the first paragraph of this section also hold.]

For a nearly uniform axial temperature distribution ($\partial T_0 / \partial x \approx 0$), Rott obtained an axial streaming velocity of

$$\langle u_2 \rangle = (1 + \alpha_1) \frac{3}{8} \frac{U_1^2}{c} \left(1 - \frac{2r^2}{R^2}\right) \sin\left(\frac{\pi x}{L}\right). \quad (4)$$

The constant α_1 represents a correction to the magnitude of the streaming velocity and is given by

$$\alpha_1 = \frac{2}{3} (1 - \beta)(\gamma - 1) \frac{\sqrt{\text{Pr}}}{1 + \text{Pr}}, \quad (5)$$

where γ is the ratio of the isobaric to isochoric specific heats and $\text{Pr} = \nu/\kappa$ is the Prandtl number. A typical value for α_1 in air ($\gamma = 1.4$, $\text{Pr} = 0.71$, and $\beta = 0.77$) is $\alpha_1 = 0.030$, which predicts an axial streaming velocity that is 3.0% larger than the result given by Eq. (2).

A result similar to Eq. (4) was later obtained independently by Qi.⁹ However, his proposed correction,

$$\alpha_1 = \frac{2}{3} (\gamma - 1) \frac{\sqrt{\text{Pr}}}{1 + \text{Pr}}, \quad (6)$$

neglects the dependence of viscosity and thermal conductivity on temperature ($\beta = 0$), predicting an axial streaming velocity that is 13.1% larger ($\alpha_1 = 0.131$) than the result given by Eq. (2).

Hamilton *et al.*¹⁴ considered analytically how heat conduction and the dependence of viscosity and thermal conductivity on temperature affect the acoustic streaming generated by standing waves in tubes of arbitrary width. For the case of a wide tube ($R \gg \delta_\nu$), their results are in agreement with the results of Rott.⁸

The influence of fluid inertia on the streaming-velocity field was studied computationally by Menguy and Gilbert.¹⁵ They concluded that the effect of fluid inertia is to distort the shape of the outer streaming cells: the location of maximum streaming velocity along the centerline ($r=0$) shifts from the midpoint ($x=L/2$) toward the velocity antinode ($x=L$), while the streaming velocity just outside the acoustic boundary layer ($r \approx \pm R$) remains unchanged. This distortion scales with the Reynolds number

$$\text{Re} = \left(\frac{U_1}{c}\right)^2 \left(\frac{R}{\delta_\nu/\sqrt{2}}\right)^2, \quad (7)$$

and is negligible for $\text{Re} \ll 1$.

III. PREVIOUS MEASUREMENTS

Only a few studies pertaining to the measurement of the streaming-velocity field in a standing-wave resonator are available in the literature. In all of these studies, optical methods were used to measure the motion of smoke particles suspended in air ($c \approx 345$ m/s).

The first known study was by Andrade,¹⁶ who photographed the smoke patterns produced by streaming in a cylindrical resonator ($2R=3.5$ cm, $\lambda=56.5$ cm, $U_1=1.3$ m/s, $Re\approx 0.5$). From these photographs, the flow streamlines were obtained and compared to the theoretical results of Rayleigh with good success.

A more quantitative investigation was carried out by Sharpe *et al.*² using the particle image velocimetry (PIV) technique. (A typical PIV system consists of a sheet of light for illuminating small tracer particles suspended in the working fluid and a photographic system for imaging the particle motion in two dimensions.) Sharpe *et al.* measured the axial and transverse components of the streaming-velocity field in the vicinity of an acoustic-velocity node ($x\approx 0$) in a cylindrical resonator ($2R=2.35$ cm, $\lambda\approx 13$ cm, $U_1\approx 2.5$ m/s, $Re\approx 1.0$). They showed the axial streaming velocity to vary parabolically with respect to the transverse coordinate, with a magnitude consistent with the measured acoustic-pressure amplitude.

Arroyo and Greated³ used stereoscopic PIV to measure all three spatial components of the streaming-velocity field in the vicinity of an acoustic-velocity node ($x\approx 0$) in a resonator of rectangular cross section (transverse dimensions= 2.22 cm \times 2.34 cm, $\lambda=18.0$ cm, $U_1\approx 1.6$ m/s, $Re\approx 0.3$). They, too, reported a parabolic variation of axial streaming velocity with respect to the transverse coordinate, with a magnitude consistent with the measured acoustic-pressure amplitude. Additionally, the precision of these measurements was better than that reported by Sharpe *et al.*²

Hann and Greated⁴ used PIV to simultaneously measure two of the spatial components of both the acoustic- and the streaming-velocity fields in the vicinity of an acoustic-velocity node ($x\approx 0$) in a resonator of square cross section (transverse dimension= 3.0 cm, $\lambda=21.4$ cm, $U_1=2.1$ m/s, $Re\approx 1.0$). Because their work focused primarily on the signal-processing technique used to separate the acoustic- and streaming-velocity fields, they did not address the spatial variation of the streaming velocity.

Rockliff *et al.*⁵ used PIV to measure the axial and transverse components of the streaming-velocity field in the vicinity of an acoustic-velocity node ($x\approx 0$) in a cylindrical resonator ($2R$ =not specified, $\lambda\approx 17$ cm, $U_1\approx 2$ m/s, Re =unknown). However, no new observations were reported.

Thompson and Atchley⁶ used the laser Doppler anemometry (LDA) technique to simultaneously measure the axial components of the acoustic- and streaming-velocity fields in a cylindrical resonator ($2R=4.65$ cm, $\lambda=1.11$ m). They reported that at low to moderate acoustic amplitudes ($U_1=0.57$ and 2.3 m/s, $Re=0.36$ and 5.7) the variation of axial streaming velocity with respect to the transverse coordinate r midway along the streaming cell ($x=L/2$) was parabolic, whereas at higher acoustic amplitudes ($U_1=4.2$ m/s, $Re=19$) the variation was not parabolic, but rather resembled a sixth-order polynomial. Although the precision of these measurements was not as good as that reported by Sharpe *et al.*² and Arroyo and Greated,³ the magnitudes of the measured streaming were consistent with the measured acoustic-velocity amplitudes.

Although several of these previous studies reported a parabolic variation of axial streaming velocity with respect to the transverse coordinate, none of them addressed the variation of axial streaming velocity with respect to the axial coordinate, the variation with respect to the acoustic amplitude at low Reynolds numbers ($Re<1$), or the corrections to the streaming magnitude proposed by Rott⁸ and Qi⁹ (see Sec. II).

IV. LDA AND BSA

The LDA technique was first applied to the measurement of acoustic fluid velocity by Taylor,¹⁷ who studied monofrequency standing waves in air. The technique was later refined and extended by Vignola *et al.*,¹⁸ who used water as a working fluid. More recently, LDA with burst spectrum analysis (BSA) has been used: by Bailliet *et al.*¹⁹ to measure acoustic velocity and power flow within a thermoacoustic resonator; by Thompson and Atchley⁶ to measure acoustic-streaming velocity in a standing-wave resonator (see Sec. III); and by Biwa *et al.*²⁰ to measure acoustic-streaming velocity and acoustic intensity associated with planar traveling waves in a cylindrical wave guide.

In a typical LDA system, two coherent laser beams intersect to form a series of interference fringes at the location where fluid velocity is to be measured (the fringe volume). When a small tracer particle suspended in the fluid passes through these fringes, the intensity of the light scattered from the particle is modulated at a frequency proportional to the spatial component of particle velocity that is normal to the fringes. It is assumed that the tracer particle follows the motion of the fluid, which Taylor¹⁷ showed to be the case for smoke in air at acoustic frequencies below approximately 5 kHz.

If one of the laser beams is frequency shifted by an amount ΔF relative to the other using an acousto-optic modulator (Bragg cell), then the intensity of the light scattered from a stationary tracer particle will be modulated at a frequency of ΔF (instead of at a frequency of zero, as in the non-frequency-shifted case). In addition to enabling the detection of stationary or low-speed particles, this frequency shift enables the determination of the sign of the velocity: a modulation frequency greater than ΔF corresponds to a particle moving in the "positive" direction, whereas a modulation frequency less than ΔF corresponds to a particle moving in the "negative" direction.

Whenever a tracer particle enters the fringe volume, a detection event known as a "burst" occurs. If the peak-to-peak particle displacements are much larger than the fringe volume, then velocity information can be extracted from these bursts using the BSA method.¹⁰ The intensity of the scattered light is measured using a photodetector, and an analog voltage is produced. This signal is bandpass filtered and frequency demodulated by $-\Delta F$ to produce both in-phase and quadrature signals that can be treated as the real and imaginary components of a single complex signal. This signal is low-pass filtered and sampled during a time interval (the BSA record interval) short relative to the acoustic period $\tau=2\pi/\omega$, such that the modulation frequency of the signal (which corresponds to the particle velocity) remains nearly

constant during that interval. The discrete Fourier transform (DFT) of the sampled signal is then calculated and the center frequency of the main spectral peak is obtained by curve fitting and interpolation. From this frequency, which can be either positive or negative, the tracer-particle velocity is obtained.²¹

V. FOURIER AVERAGING

The particle velocities obtained from a large number of bursts constitute a record of the fluid velocity at the measurement location sampled as a function of time. Analysis of this sample record is complicated by several factors: (1) The spatial distribution of tracer particles within the fluid is random, resulting in a nonuniform time interval between velocity samples. (2) The rate at which samples are acquired (the data rate) is proportional to the fluid velocity, causing the probability distribution of the samples to be skewed symmetrically away from zero in favor of high-speed samples. (3) The data rate is often less than the Nyquist rate (twice the highest acoustic frequency being sampled), resulting in acoustic-frequency aliasing. (4) The time-averaged signal component, which corresponds to the desired acoustic-streaming velocity, is typically orders of magnitude smaller than the oscillatory signal component.²⁴

The velocity waveform can be reconstructed from the sample record using the Fourier averaging (FA) method proposed by Sonnenberger *et al.*¹¹ The acoustic period τ is divided into N equal-width phase bins (indexed by bin number $n=1,2,\dots,N$) and the velocity samples are sorted by phase into the appropriate bins (e.g., all samples having an arrival time approximately equal to an integer multiple of τ are grouped together into a single phase bin at $t=\tau$; all samples having an arrival time approximately equal to an integer multiple of τ plus some value $0 < t_1 \leq \tau$ are grouped together into a single phase bin at $t=t_1$). The mean velocity within each bin, $\bar{y}[n]$, is calculated from

$$\bar{y}[n] = \frac{1}{I[n]} \sum_{i=1}^{I[n]} y[n,i], \quad (8)$$

where $y[n,i]$ is the i th sample in the n th bin, $I[n]$ is the total number of samples in the n th bin, and square brackets [...] are used to denote that the arguments are integer indices. The values $\bar{y}[n]$ represent N evenly spaced samples of the average velocity waveform during one period.

A series of complex Fourier coefficients $\tilde{Y}[z]$ (indexed by harmonic number $z=0,1,\dots,Z$) is then calculated from the phase-bin-averaged velocity data using the discrete Fourier transform

$$\tilde{Y}[z] = \frac{2}{N} \sum_{n=1}^N \bar{y}[n] e^{-j2\pi zn/N}, \quad (9)$$

for the harmonics $z=1,2,\dots,Z$ (the time-harmonic signal components, which correspond to the acoustic velocity), and

$$Y[0] = \frac{1}{N} \sum_{n=1}^N \bar{y}[n], \quad (10)$$

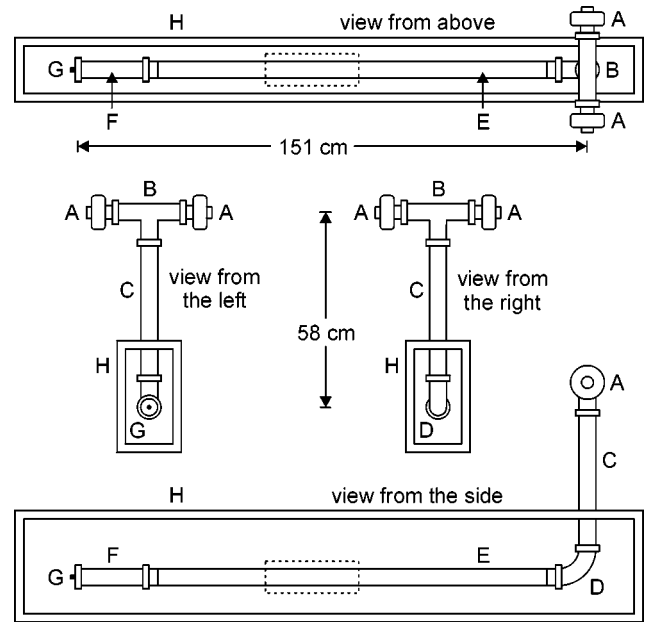


FIG. 2. The experimental apparatus. (A) Two compression drivers are connected to (B) a stainless-steel tee on top of (C) a 40 cm stainless-steel tube. This vertical section is connected by (D) a stainless-steel 90 deg elbow to a horizontal section comprised of (E) a 122 cm borosilicate-glass tube and (F) a 20 cm stainless-steel tube. The resonator is terminated by (G) a rigid stainless-steel end cap. The horizontal section is suspended inside (H) a glass-walled enclosure. The dotted box indicates the location of the streaming cell where measurements are made, which is similar to the cell shown in Fig. 1.

for $z=0$ (the time-averaged signal component, which corresponds to the streaming velocity). A smoothed velocity waveform can then be synthesized using the inverse transform

$$y(t) = \Re \left(\sum_{z=0}^{Z'} \tilde{Y}[z] e^{j\omega z t} \right), \quad (11)$$

where the Fourier series is truncated at harmonic number $Z' < Z$, below which the coefficients are considered statistically significant.²⁶

VI. PROCEDURE

The apparatus shown in Fig. 2 has been used to study the outer streaming-velocity field generated by a monofrequency acoustic standing wave in air at low Reynolds numbers ($Re < 1$). The standing-wave resonator is constructed primarily of stainless-steel tubing (type 304, inner diameter $2R=4.75$ cm, wall thickness=1.7 mm) and has a long section made of borosilicate glass (inner diameter $2R=4.65$ cm, wall thickness=2.0 mm) for making velocity measurements using an LDA/BSA system. The individual pieces are joined together using quick-connect clamps that facilitate the introduction of smoke from burned cotton rope and the cleaning of smoke residue from inside the resonator. (The diameters of the smoke particles are on the order of $1 \mu\text{m}$, measured using a Philips XL30 ESEM scanning electron microscope.)

The two compression drivers (University Sound 1828C or 1829BT) are driven sinusoidally in phase to produce a

standing wave at the fourth natural mode of the air column, which corresponds to a wavelength of $\lambda=1.11$ m and a fundamental frequency of $f=\omega/2\pi=311$ Hz when using the 1828C drivers, or $\lambda=1.13$ m and $f=308$ Hz when using the 1829BT drivers. (At these frequencies in air, the thickness of the viscous acoustic boundary layer is $\delta_v=0.13$ mm). The amplitude of the first harmonic of pressure at its antinode is monitored using a piezoresistive gauge-pressure transducer (Endevco 8510B-5) mounted flush in the rigid end cap.

The glass-walled enclosure helps to isolate the resonator from temperature gradients within the laboratory that can cause unwanted ambient convection currents. In preliminary experiments done without the enclosure, thermally induced currents were observed within the resonator that were as large in magnitude as the acoustic streaming. Elevating the drivers also helps to reduce unwanted ambient convection currents by trapping the heat generated by the drivers within the vertical section of the resonator. Several type-E thermocouples (Omega Engineering TT-E-30-SLE) are attached to the outer wall of the resonator at various locations in order to monitor the spatial distribution of mean temperature, which has a nominal value of 23 °C and is uniform within 0.1 °C along the streaming cell of interest, even in the presence of a steady-state acoustic field.

A barometric-pressure transducer (SenSym ASCX15AN) is located within the enclosure so that, in combination with the gauge-pressure transducer, the mean absolute pressure within the resonator (nominally 97 kPa) can be determined.

The axial component of streaming velocity is measured using a single-component LDA system (Dantec Measurement Technology FlowLite). The HeNe laser has an optical wavelength of 632.8 nm, a power of 10 mW, and a frequency shift of $\Delta F=40$ MHz in one of the beams. The dimensions of the fringe volume are $77 \mu\text{m}$ in the direction of measurement by $636 \mu\text{m}$ wide, and the fringe spacing is $2.63 \mu\text{m}$ (determined by measuring the intersection angle of the two beams geometrically). The LDA probe is mounted on a linear positioning system that translates in both the axial and the transverse directions in the horizontal plane. Each time a tracer particle passes through the LDA fringe volume, its velocity is sampled and the corresponding arrival time and transit time are recorded by a BSA system (Dantec Measurement Technology 57N11) connected to the LDA photomultiplier output. The absolute accuracy of the velocity measurements is approximately 0.3%, based on the uncertainty in the intersection angle of the two LDA laser beams. However, the random errors in the individual velocity samples are much larger than 0.3% and vary with the BSA settings used.

A schematic diagram of the instrumentation is shown in Fig. 3. The BSA system, the sine-wave generator, and the two data-acquisition systems are synchronized to prevent any timing jitter or drift from interfering with the phase-bin-averaging procedure (see Sec. V). In preliminary measurements, it was observed that not having the clocks synchronized resulted in cumulative errors in the phases of the velocity samples, causing the samples to be incorrectly sorted into the wrong phase bins.

In preparation for making measurements, the electronics

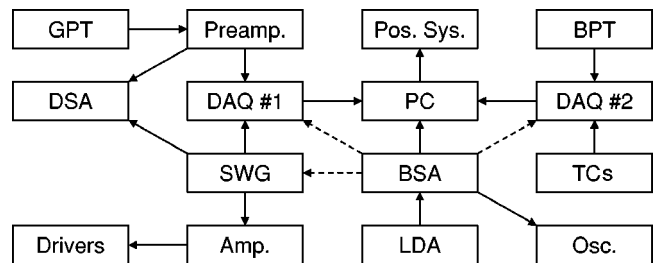


FIG. 3. Schematic diagram of the instrumentation. The sine-wave generator (SWG) is connected to the two compression drivers through an amplifier (Amp.). The signal from the gauge-pressure transducer (GPT) flows through a preamplifier (Preamp.) before being sampled by a data-acquisition system (DAQ #1). The signals from the GPT preamplifier and the SWG are also monitored using a dynamic signal analyzer (DSA). The signal from the LDA system is processed by the BSA system. The envelope of the LDA signal is also monitored using an oscilloscope (Osc.). The barometric pressure transducer (BPT) and thermocouples (TCs) are connected to another data-acquisition system (DAQ #2). The BSA and the two DAQs are connected to a personal computer (PC) that also controls the positioning system (Pos. Sys.) for the LDA probe. The BSA, SWG, and the two DAQs are synchronized (dashed lines).

are allowed to warm up, the resonator is allowed to reach a uniform temperature distribution, and the rigid end cap and one of the compression drivers are temporarily removed so that smoke from burned cotton rope can be introduced into the resonator. Once a sufficient smoke density has been achieved at the location where measurements are to be made, the end cap and the compression driver are replaced and the air within the resonator is allowed to come to rest. The acoustic field is then switched on and measurements of velocity, gauge pressure, barometric pressure, and temperature are made every few minutes. Between measurements, the standing wave is maintained at a constant wavelength and amplitude by adjusting the frequency and amplitude produced by the sine-wave generator.

VII. COEFFICIENT UNCERTAINTIES

The sample records of velocity contain between 1000 and 30 000 samples acquired during a time interval of 10 to 115 s. For each record, the FA method¹¹ (see Sec. V) is used to calculate a series of complex Fourier coefficients representing the time-averaged ($z=0$) and time-harmonic ($z=1,2,\dots,Z$) signal components.

The uncertainties in the magnitudes and phases of these Fourier coefficients can be calculated based on the scatter of the samples within each phase bin. Propagation-of-error analysis allows the uncertainty in a function $g(\chi_1, \chi_2, \dots, \chi_N)$ with uncorrelated arguments $\chi_1, \chi_2, \dots, \chi_N$ to be computed from the uncertainties in the arguments, $\sigma_{\chi_1}, \sigma_{\chi_2}, \dots, \sigma_{\chi_N}$, according to the formula

$$\sigma_g = \left\{ \sum_{n=1}^N \left(\frac{\partial g}{\partial \chi_n} \sigma_{\chi_n} \right)^2 \right\}^{1/2}. \quad (12)$$

The function $g(\chi_1, \chi_2, \dots, \chi_N)$ represents the magnitude or phase of a Fourier coefficient calculated from Eq. (9) or Eq. (10), and the arguments $\chi_1, \chi_2, \dots, \chi_N$ represent the mean bin velocities $\bar{y}[n]$ calculated from Eq. (8). The uncertainties in the mean bin velocities are first calculated using the

equation for the standard error of the mean (SEM)

$$\sigma_{\bar{y}}[n] = \frac{1}{\sqrt{I[n]}} \left\{ \sum_{i=1}^{I[n]} \frac{(y[n,i] - \bar{y}[n])^2}{I[n] - 1} \right\}^{1/2}. \quad (13)$$

A. Uncertainty in the acoustic-velocity magnitude

The form of Eq. (12) used to compute the uncertainty in $Y[z]$, the magnitude of $\tilde{Y}[z]$, is

$$\sigma_Y[z] = \left\{ \sum_{n=1}^N \left(\frac{\partial Y[z]}{\partial \bar{y}[n]} \sigma_{\bar{y}}[n] \right)^2 \right\}^{1/2}, \quad (14)$$

for $z = 1, 2, \dots, Z$, where $\sigma_{\bar{y}}[n]$ is given by Eq. (13). Using the relations

$$Y[z] = \{\tilde{Y}[z](\tilde{Y}[z])^*\}^{1/2}, \quad (15)$$

and

$$\tilde{Y}[z] = Y[z]e^{j\phi[z]}, \quad (16)$$

where $\phi[z]$ is the phase of $\tilde{Y}[z]$ and $(\tilde{Y}[z])^*$ is the complex conjugate of $\tilde{Y}[z]$, the partial derivative in Eq. (14) is written as

$$\frac{\partial Y[z]}{\partial \bar{y}[n]} = \frac{1}{2} \left\{ e^{j\phi[z]} \frac{\partial (\tilde{Y}[z])^*}{\partial \bar{y}[n]} + e^{-j\phi[z]} \frac{\partial \tilde{Y}[z]}{\partial \bar{y}[n]} \right\}. \quad (17)$$

Substituting the expression for $\tilde{Y}[z]$ given by Eq. (9) into Eq. (17) yields

$$\frac{\partial Y[z]}{\partial \bar{y}[n]} = \frac{2}{N} \cos\left(2\pi z \frac{n}{N} + \phi[z]\right). \quad (18)$$

When this result is substituted into Eq. (14), the uncertainty in $Y[z]$ is given by

$$\sigma_Y[z] = \left(\frac{4}{N^2} \sum_{n=1}^N \left\{ \sigma_{\bar{y}}[n] \cos\left(2\pi z \frac{n}{N} + \phi[z]\right) \right\}^2 \right)^{1/2}, \quad (19)$$

for $z = 1, 2, \dots, Z$.

B. Uncertainty in the streaming velocity

The form of Eq. (12) used to compute the uncertainty in $Y[0]$ is given by Eq. (14) for $z=0$, where the partial derivative is equal to

$$\frac{\partial Y[0]}{\partial \bar{y}[n]} = \frac{1}{N}, \quad (20)$$

from Eq. (10), and $\sigma_{\bar{y}}[n]$ is given by Eq. (13). When these results are combined, the uncertainty in $Y[0]$ is given by

$$\sigma_Y[0] = \left\{ \frac{1}{N^2} \sum_{n=1}^N (\sigma_{\bar{y}}[n])^2 \right\}^{1/2}. \quad (21)$$

C. Uncertainty in the acoustic-velocity phase

The form of Eq. (12) used to compute the uncertainty in $\phi[z]$, the phase of $\tilde{Y}[z]$, is

$$\sigma_{\phi}[z] = \left\{ \sum_{n=1}^N \left(\frac{\partial \phi[z]}{\partial \bar{y}[n]} \sigma_{\bar{y}}[n] \right)^2 \right\}^{1/2}, \quad (22)$$

for $z = 1, 2, \dots, Z$, where $\sigma_{\bar{y}}[n]$ is given by Eq. (13). An expression containing the partial derivative in Eq. (22) is obtained by equating Eqs. (9) and (16) and taking the partial derivative with respect to $\bar{y}[n]$

$$\frac{2}{N} e^{-j2\pi zn/N} = e^{j\phi[z]} \frac{\partial Y[z]}{\partial \bar{y}[n]} + jY[z] e^{j\phi[z]} \frac{\partial \phi[z]}{\partial \bar{y}[n]}. \quad (23)$$

Substituting the expression for $\partial Y[z]/\partial \bar{y}[n]$ given by Eq. (18) into Eq. (23) and solving for $\partial \phi[z]/\partial \bar{y}[n]$ yields

$$\frac{\partial \phi[z]}{\partial \bar{y}[n]} = -\frac{1}{Y[z]} \frac{2}{N} \sin\left(2\pi z \frac{n}{N} + \phi[z]\right). \quad (24)$$

When this result is substituted into Eq. (22), the uncertainty in $\phi[z]$ is given by

$$\sigma_{\phi}[z] = \frac{1}{Y[z]} \left(\frac{4}{N^2} \sum_{n=1}^N \left\{ \sigma_{\bar{y}}[n] \sin\left(2\pi z \frac{n}{N} + \phi[z]\right) \right\}^2 \right)^{1/2}, \quad (25)$$

for $z = 1, 2, \dots, Z$.

D. Discussion

In a canonical implementation of the FA method, all of the coefficients lying above a fixed cutoff frequency would be discarded. In the present study, however, the statistical significance of each coefficient is determined based on whether its magnitude is greater than the 4σ uncertainty in its magnitude, $Y[z] > 4\sigma_Y[z]$ (which corresponds to the 99.99% confidence interval). Insignificant coefficients are discarded, and the significant coefficients are used to synthesize a smoothed velocity waveform [compare to Eq. (11)]

$$y(t) = \Re \left(\sum_{z'} \tilde{Y}[z'] e^{j\omega z' t} \right), \quad (26)$$

where the index z' indicates that the sum includes only the statistically significant coefficients.

Although this method is designed to process BSA velocity data that exhibit a nonuniform time between samples and a skewed sampling distribution, it has also been successfully used to process data that do not exhibit these shortcomings, such as would be sampled using a standard data acquisition unit (provided that the sample rate is not an integer multiple of the acoustic frequency).

VIII. LAGRANGIAN STREAMING VELOCITY

The velocity of the fluid sampled at a fixed position in space (the fringe volume) using the LDA technique is the Eulerian fluid velocity. However, the uncertainties [see Eq. (21)] in the Eulerian streaming velocities determined using the FA method¹¹ [see Eq. (10)] are unacceptably large relative to the magnitudes of the streaming predicted by theory. (In the present study, the 2σ uncertainties are on the order of ± 0.3 mm/s and the streaming velocities are on the order of 0.7 mm/s.) These errors can be reduced somewhat by acquiring and averaging over a greater number of samples.

More troubling, however, are bias errors observed in the sampled velocity signal. These errors, which are on the order of 0.4 mm/s or 0.1% of the amplitude of the fluid velocity

being measured, would be negligible under most circumstances. They do become significant, however, when measuring very small streaming velocities in the presence of a much larger acoustic field, as in the present study. The mechanism responsible for these errors is not well understood, but the source of these errors appears to be the LDA/BSA system itself.

Because of these difficulties, a more satisfying comparison to theory is made using the Lagrangian streaming velocity—defined as the time-averaged velocity of an individual tracer particle in the vicinity of the fringe volume.

A. Comparison of Eulerian and Lagrangian velocities

The relationship between the Eulerian and Lagrangian velocities in a one-dimensional standing wave is here considered. In the derivation that follows, **bold** typeface is used to denote vector quantities.

Let \mathbf{u} represent the Eulerian fluid velocity at a fixed position in space and \mathbf{u}_L represent the Lagrangian velocity of a particle moving with the fluid near that position. Furthermore, let \mathbf{u} and \mathbf{u}_L be expandable in terms of the acoustic Mach number M

$$\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2 + O(M^3), \quad (27)$$

$$\mathbf{u}_L = \mathbf{u}_{L1} + \mathbf{u}_{L2} + O(M^3), \quad (28)$$

where \mathbf{u}_1 and \mathbf{u}_{L1} are approximations valid to first order in M , \mathbf{u}_2 and \mathbf{u}_{L2} are second-order corrections to the first-order quantities, and $O(M^3)$ represents third- and higher-order corrections. The relationship between \mathbf{u} and \mathbf{u}_L valid to $O(M^2)$ is given by Nyborg²⁸ as

$$\mathbf{u}_{L1} + \mathbf{u}_{L2} = \mathbf{u}_1 + \mathbf{u}_2 + \left[\left(\int \mathbf{u}_1 dt \right) \cdot \nabla \right] \mathbf{u}_1. \quad (29)$$

For the one-dimensional acoustic standing-wave field given by [compare to Eq. (1)]

$$\mathbf{u}_1 = \Re(\tilde{U}_1 e^{j\omega t}) \sin\left(\frac{2\pi x}{\lambda}\right) \hat{\mathbf{x}}, \quad (30)$$

where $\hat{\mathbf{x}}$ is the unit vector in the $+x$ direction, Eq. (29) takes the form

$$\mathbf{u}_{L1} + \mathbf{u}_{L2} = \mathbf{u}_1 + \mathbf{u}_2 + \frac{1}{4c} \Im(\tilde{U}_1^2 e^{2j\omega t}) \sin\left(\frac{4\pi x}{\lambda}\right) \hat{\mathbf{x}}, \quad (31)$$

where \Im denotes taking the imaginary part of the complex argument. The first-order velocities are shown to be identical by limiting Eq. (31) to $O(M)$

$$\mathbf{u}_{L1} = \mathbf{u}_1. \quad (32)$$

The second-order streaming velocities are shown to be identical by taking the average of Eq. (31) with respect to time

$$\langle \mathbf{u}_{L2} \rangle = \langle \mathbf{u}_2 \rangle. \quad (33)$$

The second-order acoustic velocities differ only by the third term on the right side of Eq. (31), which represents a correction to the magnitude and phase of the second harmonic of velocity. For an acoustic-velocity amplitude U_1 on the order of 1 m/s (as in the present study), the magnitude of this second-harmonic correction is less than 0.08% of the magni-

tude of the first harmonic. Including such a small correction is unnecessary because it would result in a negligible change in the amplitude and shape of the velocity waveform.

B. Determination of Lagrangian streaming velocity

Figures 4(a) and (b) show samples of the axial component of the Eulerian velocity obtained from a single tracer particle. The particle oscillates axially due to the acoustic standing wave and is being convected in the positive axial direction by the streaming. A sample is usually recorded each time the particle passes through the fringe volume, which occurs twice per acoustic cycle.

An approximate expression for the particle's time-dependent axial position relative to the fringe volume is obtained by integrating Eq. (26) with respect to t

$$\xi(t) = \xi_0 + Y[0]t + \Re\left(\sum_{z'} \frac{\tilde{Y}[z']}{j\omega z'} e^{j\omega z' t}\right), \quad (34)$$

where ξ_0 is a constant of integration determined by the initial position of the particle and the index z' indicates that the sum includes only the statistically significant coefficients, except for $z'=0$, which is treated separately in the term $Y[0]t$. If the value of $Y[0]$ calculated using the FA method [see Eq. (10)] is used in Eq. (34), then the samples of the relative particle position—obtained by evaluating Eq. (34) for values of t equal to the arrival times of the velocity samples, $t[n,i]$ —will lie along a line having a nonzero slope, as shown in Fig. 5. This erroneous result is due neither to physical causes nor to the FA method, but rather to an instrumentation error in the LDA/BSA system. The correct value of $Y[0]$, in which this bias error is absent, is obtained by adjusting $Y[0]$ iteratively until the relative-position samples lie instead along a horizontal line, as shown in Figs. 4(c) and (d). The value of $Y[0]$ determined in this manner is the particle's Lagrangian streaming velocity.²⁹

An interpretation of the elliptical pattern formed by the velocity samples in Fig. 4(a) is now easily given. When the tracer particle first encounters the fringe volume at time $t = 6.45$ s, its observed speed is zero because the particle's equilibrium position lies in the $-x$ direction a distance equal to the peak acoustic displacement [$(U_1/\sqrt{2})/\omega = 0.30$ mm], and the particle is being sampled at the extreme of its motion. At time $t = 6.85$ s, the observed speed of the particle has increased to its maximum value because the particle has been convected a distance 0.30 mm in the $+x$ direction and its equilibrium position is centered on the fringe volume. At time $t = 7.26$ s, the observed speed of the particle has decreased to zero because the particle has been convected an additional 0.30 mm in the $+x$ direction and is once again being sampled at the extreme of its motion. The particle's Lagrangian streaming velocity is equal to the total distance convected divided by the corresponding elapsed time: $(0.60 \text{ mm})/(0.81 \text{ s}) = 0.74 \text{ mm/s}$.

If the particle motion is entirely in the axial direction, as in the example considered here, then the elliptical pattern formed by the velocity samples in Fig. 4(a) will be closed and the horizontal line segment formed by the relative-

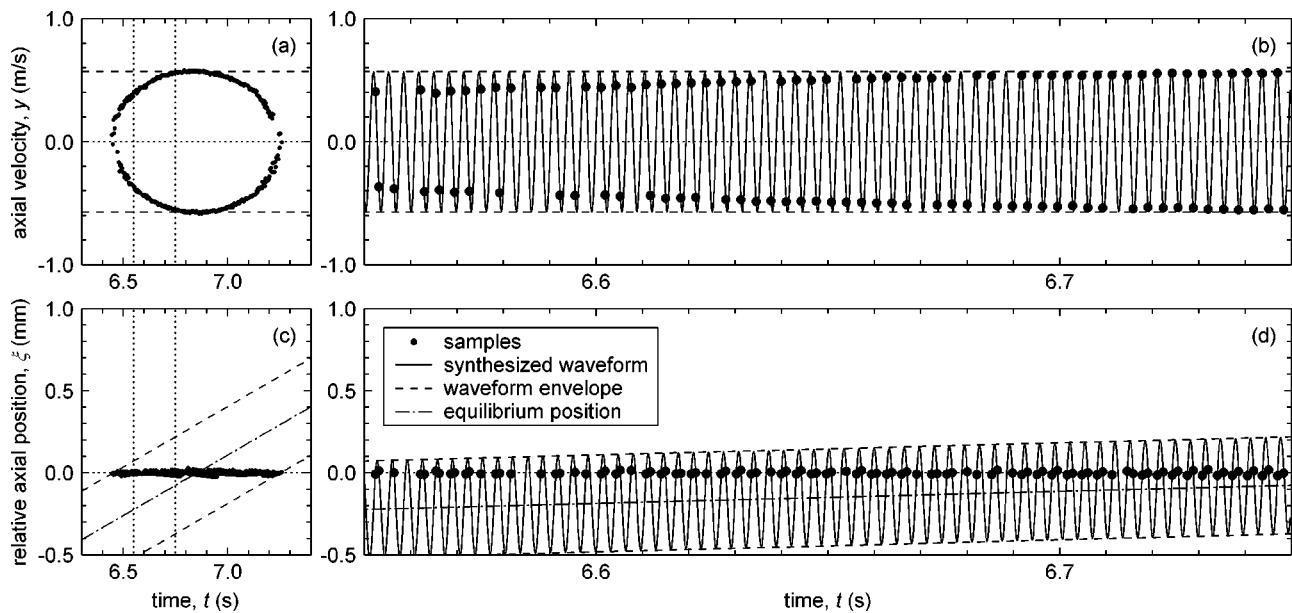


FIG. 4. Panel (a) shows samples of the axial component of the Eulerian velocity obtained from a single tracer particle oscillating near the center of the streaming cell ($x=L/2, r=0$). A detailed view of the area between the two vertical dotted lines is shown in panel (b). The solid line in (b) represents the velocity waveform calculated from Eq. (26) and the dashed lines show the waveform envelope ($U_1/\sqrt{2}=0.57$ m/s, $Re=0.36$). Panels (c) and (d) show corresponding samples of the particle's axial position relative to the fringe volume, calculated from Eq. (34). The solid line in (d) represents the relative-position waveform, also calculated from Eq. (34), and the dashed lines show the waveform envelope. The dot-dashed line represents the time-dependent equilibrium position of the particle as it is convected in the $+x$ direction by the streaming. The slope of that line (0.74 mm/s) is the particle's Lagrangian streaming velocity.

position samples in Fig. 4(c) will span the dashed lines. However, such will not be the case for a particle observed at a position where its motion also has a component in the transverse direction due to the streaming. Even so, the method presented here will still accurately determine the particle's Lagrangian streaming velocity.

IX. MEASURED ONSET OF STREAMING

In practice, each sample record contains observations of multiple tracer particles. Although the elliptical patterns formed by the different particles often overlap in time, it is

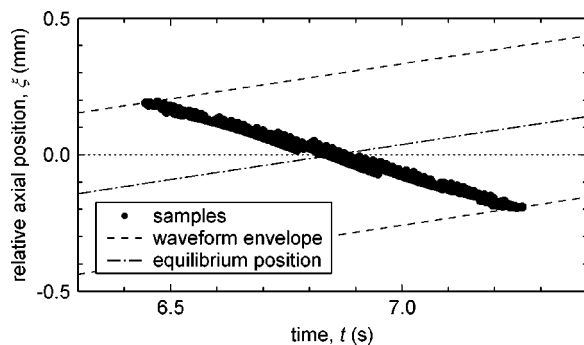


FIG. 5. Incorrect determination of Lagrangian streaming velocity. The value of $Y[0]$ calculated using the FA method [see Eq. (10)] contains a bias error produced by the LDA/BSA system. When this value of $Y[0]$ is used in Eq. (34) to obtain samples of the relative particle position, these samples lie along a line having a nonzero slope. This result implies incorrectly that the position where the particle is sampled (the LDA fringe volume) is not stationary, but moves instead at a constant speed in the negative direction. By contrast, the Lagrangian streaming velocity in Figs. 4(c) and (d) is determined correctly: the relative-position samples lie instead along a horizontal line, consistent with a stationary fringe volume.

not difficult to determine the Lagrangian streaming velocity of each particle using the method presented in Sec. VIII.

Figure 6 shows a measurement of the onset of streaming associated with the acoustic field being switched on. Each

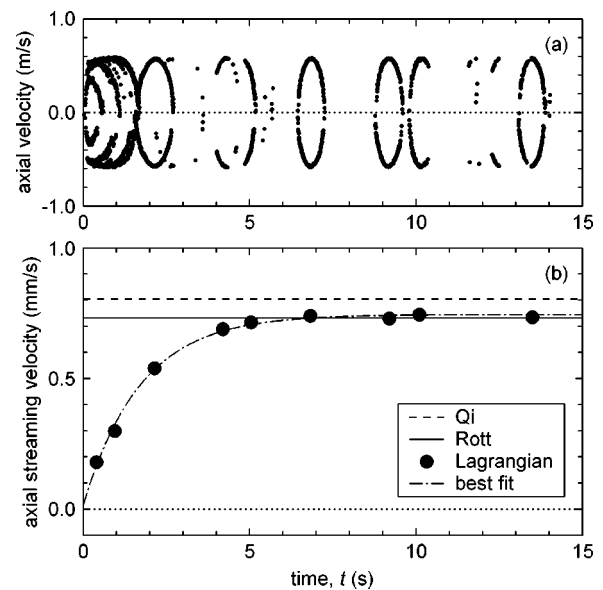


FIG. 6. Onset of streaming measured at the center of the streaming cell ($x=L/2, r=0$) as a function of time after the acoustic field is switched on ($t=0$ s). In panel (a), samples of the axial component of fluid velocity are shown. The acoustic velocity increases rapidly from zero to its steady-state value ($U_1/\sqrt{2}=0.57$ m/s, $Re=0.36$) within approximately 0.3 s. In (b), the axial component of Lagrangian streaming velocity determined from the samples in (a) increases exponentially from zero to 95% of its steady-state value within approximately 5.2 s. The theoretical steady-state streaming velocities are calculated from Eq. (4) using $\alpha_1=0.030$ for Rott [see Eq. (5)] and $\alpha_1=0.131$ for Qi [see Eq. (6)].

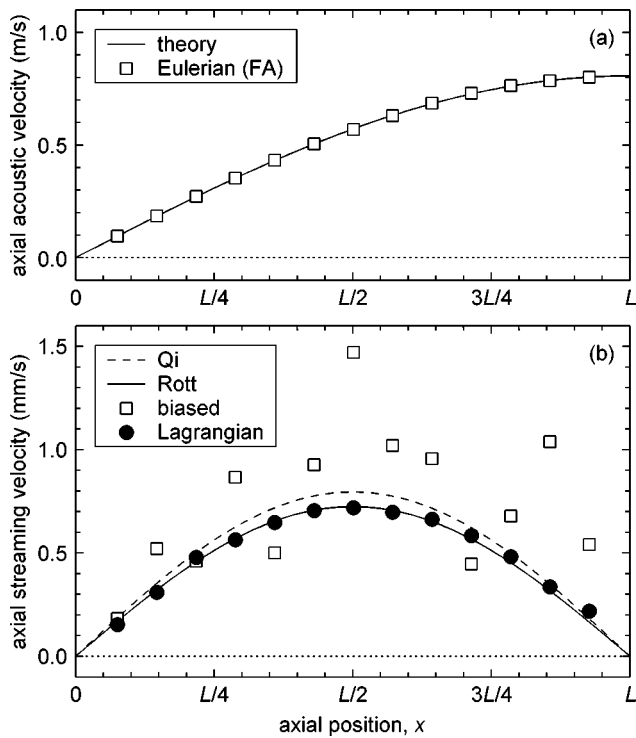


FIG. 7. Variation of the steady-state axial components of acoustic velocity and streaming velocity with respect to the axial coordinate x ($r=0$, $U_1 = 0.81$ m/s, $Re=0.36$). Panel (a) shows the amplitudes of the first harmonic of the Eulerian velocity determined using the FA method [see Eq. (9)] compared to a least-squares fit to the theory given by Eq. (1). Panel (b) shows the corresponding biased Eulerian streaming velocities determined using the FA method [see Eq. (10)] and the Lagrangian streaming velocities determined using the method presented in Sec. VIII compared to the theoretical streaming velocities calculated from Eq. (4) using $\alpha_1=0.030$ for Rott [see Eq. (5)] and $\alpha_1=0.131$ for Qi [see Eq. (6)].

value of the measured Lagrangian streaming velocity shown in Fig. 6(b) corresponds to a different tracer particle passing through the fringe volume, as indicated by the elliptical patterns shown in Fig. 6(a). (The particle at $t=6.8$ s is considered in detail in Fig. 4.) The steady-state value of the measured streaming velocity in Fig. 6(b)—given by the value of the curve fit for $t > 15$ s—corresponds to a correction to the magnitude of the streaming-velocity field of $\alpha_1=0.048$ in Eq. (4). Because the fluid is initially at rest (at $t=0$ s), it can be concluded that the ambient convection currents within the resonator are negligibly small at this location.

X. MEASURED STEADY-STATE STREAMING

When measuring the streaming velocity under steady-state conditions, all tracer particles in a given sample record have approximately the same Lagrangian streaming velocity. Therefore, an average value of the Lagrangian streaming velocity is determined that best represents the entire sample record.

Figure 7 shows the variation of the steady-state acoustic and streaming velocities with respect to the axial coordinate x . A least-squares fit between the measured values of the first harmonic of velocity and Eq. (1) is used to determine the values of the amplitude of the acoustic velocity at its antinode, U_1 , and the wavelength λ . A least-squares fit between the measured values of Lagrangian streaming velocity and

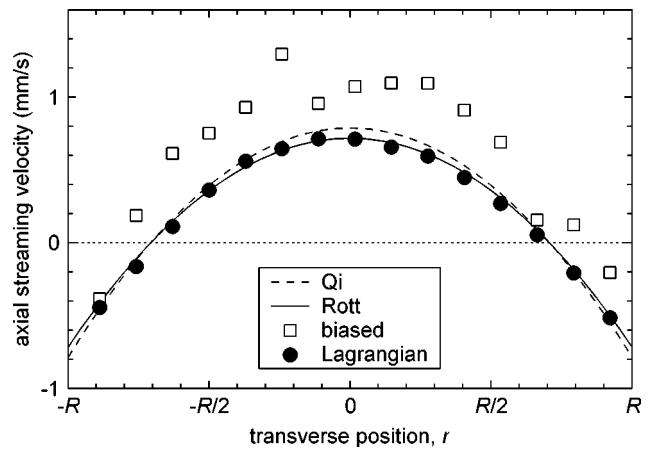


FIG. 8. Variation of the steady-state axial component of streaming velocity with respect to the transverse coordinate r ($x=L/2$, $U_1/\sqrt{2}=0.57$ m/s, $Re=0.36$). The biased Eulerian streaming velocities determined using the FA method [see Eq. (10)] and the Lagrangian streaming velocities determined using the method presented in Sec. VIII are compared to the theoretical streaming velocities calculated from Eq. (4) using $\alpha_1=0.030$ for Rott [see Eq. (5)] and $\alpha_1=0.131$ for Qi [see Eq. (6)].

Eq. (4) yields a correction to the magnitude of the streaming-velocity field of $\alpha_1=0.039$. The measured values of Eulerian streaming velocity determined using the FA method¹¹ exhibit large errors relative to the corresponding Lagrangian streaming velocities.

Figure 8 shows³⁰ the variation of the steady-state streaming velocity with respect to the transverse coordinate r . A least-squares fit between the measured values of Lagrangian streaming velocity and Eq. (4) yields a correction to the magnitude of the streaming-velocity field of $\alpha_1=0.031$.

The measured values of Eulerian streaming velocity determined using the FA method exhibit a large positive offset relative to the corresponding Lagrangian streaming velocities. When measurements of the streaming velocity are repeated with the polarity of the LDA probe reversed (by rotating the probe by 180 deg), the magnitude of the offset is approximately the same but the sign of the offset is reversed. This indicates that the offset is due neither to physical causes (see Sec. VIII A) nor to the FA method, but rather to the LDA/BSA system. The magnitude of the offset seems to vary slowly with time and to depend on the specific BSA settings used to acquire the data (such as the Doppler-frequency bandwidth, number of samples per burst, etc.).

Figure 9 shows the variation of the steady-state streaming velocity with respect to the amplitude of the acoustic velocity at its antinode, U_1 . A least-squares fit between the measured values of Lagrangian streaming velocity and Eq. (4) yields a correction to the magnitude of the streaming-velocity field of $\alpha_1=0.027$. The measured values of Eulerian streaming velocity determined using the FA method once again exhibit a large positive offset relative to the corresponding Lagrangian streaming velocities. As the BSA Doppler-frequency bandwidth is increased to accommodate larger acoustic amplitudes, the magnitude of the offset increases.

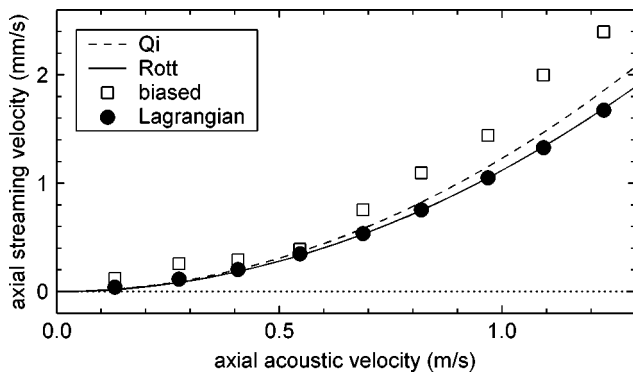


FIG. 9. Variation of the steady-state axial component of streaming velocity with respect to the amplitude of the acoustic velocity at its antinode, $U_1 < 1.3$ m/s ($Re < 0.94$). The amplitude of the acoustic velocity at its antinode ($x=L$) is determined by multiplying the acoustic-velocity amplitude at this location ($x=L/2$, $r=0$) by $\sqrt{2}$. The biased Eulerian streaming velocities determined using the FA method [see Eq. (10)] and the Lagrangian streaming velocities determined using the method presented in Sec. VIII are compared to the theoretical streaming velocities calculated from Eq. (4) using $\alpha_1=0.030$ for Rott [see Eq. (5)] and $\alpha_1=0.131$ for Qi [see Eq. (6)].

XI. CONCLUSIONS

The LDA/BSA technique (see Secs. IV and VI), when used in conjunction with the FA method¹¹ (see Sec. V) and the method presented in Sec. VIII for the determination of Lagrangian streaming velocity, is well suited to the simultaneous measurement of acoustic velocity and Lagrangian streaming velocity in a periodic acoustic field. This technique is capable of making high-precision measurements of velocity (see Sec. X) and of determining the variation of Lagrangian streaming velocity with respect to time (see Sec. IX). However, while the PIV technique (see Sec. III) is capable of collecting data simultaneously within a two-dimensional region, the LDA/BSA technique is limited to collecting data one position at a time.

The bias errors observed in the measured Eulerian streaming velocities (see Sec. X) are due neither to physical causes (see Sec. VIII A) nor to the FA method, but rather to the LDA/BSA system. This type of instrumentation error becomes inconsequential, however, when determining the Lagrangian streaming velocity using the method presented in Sec. VIII—only the measured acoustic velocity and the arrival times of the velocity samples are used; the time-averaged component of the velocity signal is irrelevant.

Equations (19), (21), and (25) (see Sec. VII), which are used to determine the uncertainties in the Fourier coefficients, represent an improvement to the FA method.¹¹ Whereas the FA method assumes a single cutoff frequency below which all of the coefficients are statistically significant, these equations permit the determination of the significance of each coefficient individually. The uncertainties computed in this manner are in agreement with the observed repeatability of the measurements.

In the previous experimental studies of acoustic streaming (see Sec. III) as well as in the present study, various smoothing methods have been used to reduce random fluctuations in the streaming measurements. Because of this, it is not straightforward to quantitatively compare the precisions of the LDA and PIV measurement techniques used. Nonethe-

less, the precisions achieved in the present study appear to be better than the precisions achieved previously.

All of the measurements of the axial component of streaming velocity presented in Secs. IX and X are consistent with the theory given by Rott⁸ (see Sec. II): the variation of axial streaming velocity with respect to the axial coordinate is sinusoidal, the variation with respect to the transverse coordinate is parabolic, and the variation with respect to the acoustic-velocity amplitude is quadratic. The corrections to the magnitude of the streaming-velocity field [see Eq. (4)] determined from the curves fit to the measured data range from $\alpha_1=0.027$ to 0.048. These values show better agreement with Rott's predicted correction of $\alpha_1=0.030$ [see Eq. (5)] than with Qi's predicted correction⁹ of $\alpha_1=0.131$ [see Eq. (6)], indicating that the magnitude of the streaming is noticeably influenced by the dependence of viscosity and thermal conductivity on temperature.

Although these results are limited to the case of a relatively low-amplitude acoustic standing wave ($Re < 1$) in the presence of a spatially uniform temperature distribution, the methods employed in the present study are also expected to be of use in the investigation of three other cases of interest: (1) the influence of fluid inertia on acoustic streaming at high Reynolds numbers ($Re > 1$), (2) the behavior of acoustic streaming in the presence of a temperature gradient, and (3) the acoustic streaming associated with a low-amplitude traveling wave in a guide.

ACKNOWLEDGMENTS

This work has been supported by the Office of Naval Research and The Pennsylvania State University Graduate Program in Acoustics and Applied Research Laboratory Exploratory and Foundational Program. The authors would like to thank: Scott Kralik and Michael Maccarone for measuring the diameters of the smoke particles; and Scott Backhaus, Tetsushi Biwa, Said Boluriaan, Murray Campbell, Andrew Doller, Steven Garrett, Joël Gilbert, Mark Hamilton, Yuri Ilinskii, Robert Keolian, Philip Morris, Matthew Poese, Victor Sparrow, Greg Swift, and Ray Wakeland for numerous helpful discussions during the course of this work.

¹J. W. Strutt, Baron Rayleigh, *The Theory of Sound*, 2nd ed. (Dover, New York, 1945), Vol. 2, Chap. XIX, Sec. 352, pp. 333–342.

²J. P. Sharpe, C. A. Greated, C. Gray, and D. M. Campbell, "The measurement of acoustic streaming using particle image velocimetry," *Acustica* **68**, 168–172 (1989).

³M. P. Arroyo and C. A. Greated, "Stereoscopic particle image velocimetry," *Meas. Sci. Technol.* **2**, 1181–1186 (1991).

⁴D. B. Hann and C. A. Greated, "The measurement of flow velocity and acoustic particle velocity using particle-image velocimetry," *Meas. Sci. Technol.* **8**, 1517–1522 (1997).

⁵M. Campbell, J. A. Cosgrove, C. A. Greated, S. Jack, and D. Rockliff, "Review of LDA and PIV applied to the measurement of sound and acoustic streaming," *Opt. Laser Technol.* **32**, 629–639 (2000).

⁶M. W. Thompson and A. A. Atchley, "Measurements of Rayleigh streaming in high-amplitude standing waves," in *Nonlinear Acoustics at the Beginning of the 21st Century*, edited by O. V. Rudenko and O. A. Sapozhnikov (MSU Faculty of Physics, Moscow, 2002), Vol. 1, pp. 183–190.

⁷M. W. Thompson and A. A. Atchley, "Measurement of acoustic streaming using laser Doppler anemometry," *Proceedings of the 18th International Congress on Acoustics*, Kyoto, Japan, 4–9 April 2004.

- ⁸N. Rott, "The influence of heat conduction on acoustic streaming," *Z. Angew. Math. Phys.* **25**, 417–421 (1974).
- ⁹Q. Qi, "The effect of compressibility on acoustic streaming near a rigid boundary for a plane traveling wave," *J. Acoust. Soc. Am.* **94**, 1090–1098 (1993).
- ¹⁰Dantec Measurement Technology A/S, P.O. Box 121, Tonsbakken 18, DK-2740 Skovlunde, Denmark, *BSA Installation & User's Guide*, 2nd ed. (Dantec, Skovlunde, Denmark, 1996), pp. 93–111.
- ¹¹R. Sonnenberger, K. Graichen, and P. Erk, "Fourier averaging: a phase-averaging method for periodic flow," *Exp. Fluids* **28**, 217–224 (2000).
- ¹²K. Schuster and W. Matz, "Über stationäre Strömungen im Kundtschen Rohr" ("On the steady flows in a Kundt's tube"), *Akust. Z.* **5**, 349–352 (1940).
- ¹³Y. Ilinskii, "Acoustic streaming in standing waves in a cylindrical tube" (2001, private communication). The sign of the transverse component of streaming velocity given by Eq. (13) in Ref. 12 is incorrect.
- ¹⁴M. F. Hamilton, Y. A. Ilinskii, and E. A. Zabolotskaya, "Thermal effects on acoustic streaming in standing waves," *J. Acoust. Soc. Am.* **114**, 3092–3101 (2003).
- ¹⁵L. Menguy and J. Gilbert, "Non-linear acoustic streaming accompanying a plane stationary wave in a guide," *Acust. Acta Acust.* **86**, 249–259 (2000).
- ¹⁶E. N. da C. Andrade, "On the circulations caused by the vibration of air in a tube," *Proc. R. Soc. London, Ser. A* **134**, 445–470 (1931).
- ¹⁷K. J. Taylor, "Absolute measurement of acoustic particle velocity," *J. Acoust. Soc. Am.* **59**, 691–694 (1976).
- ¹⁸J. F. Vignola, Y. H. Berthelot, and J. Jarzynski, "Laser detection of sound," *J. Acoust. Soc. Am.* **90**, 1275–1286 (1991).
- ¹⁹H. Bailliet, P. Lotton, M. Bruneau, V. Gusev, J. C. Valière, and B. Gzenzel, "Acoustic power flow measurement in a thermoacoustic resonator by means of laser Doppler anemometry (LDA) and microphonic measurement," *Appl. Acoust.* **60**, 1–11 (2000).
- ²⁰T. Biwa, M. W. Thompson, and A. A. Atchley, "Optical measurement of traveling-wave acoustic streaming in a wave guide," *Proceedings of the 18th International Congress on Acoustics, Kyoto, Japan, 4–9 April 2004*.
- ²¹If the peak-to-peak particle displacements are on the order of the fringe volume, then more complex signal processing may be required, as discussed by Hann and Greated (Ref. 22) and Valière *et al.* (Ref. 23).
- ²²D. Hann and C. A. Greated, "Acoustic measurements in flows using photon correlation spectroscopy," *Meas. Sci. Technol.* **5**, 157–164 (1994).
- ²³J. C. Valière, P. Herzog, V. Valeau, and G. Tournois, "Acoustic velocity measurements in the air by means of laser Doppler velocimetry: Dynamic and frequency range limitations and signal processing improvements," *J. Sound Vib.* **229**, 607–626 (2000).
- ²⁴If the peak-to-peak particle displacements are much smaller than the fringe spacing—as occurs in the measurement of acoustic streaming generated in an ultrasound beam [see, for example, the work by Duck *et al.* (Ref. 25)]—then the LDA/BSA system is insensitive to the acoustic velocity, and the streaming velocity is measured directly, which greatly simplifies the signal-processing requirements.
- ²⁵F. A. Duck, S. A. MacGregor, and D. Greenwell, "Measurement of streaming velocities in medical ultrasonic beams using laser anemometry," in *Advances in Nonlinear Acoustics*, edited by H. Hobæk (World Scientific, Singapore, 1993), pp. 607–612.
- ²⁶Thompson and Atchley showed in a previous analysis (Ref. 7) that the addition of transit-time (or residence-time) weighting (Ref. 27) to the FA method does not improve the agreement between the observed streaming velocities and the theory of Rott (Ref. 8) (see Sec. II). Additionally, they showed that if the time-equivalent phase-bin width is set equal to the BSA record interval (see Sec. IV), then their proposed recursive algorithm (Ref. 6) performs no better than the FA method, on which it is based.
- ²⁷P. Buchhave, W. K. George, Jr., and J. L. Lumley, "The measurement of turbulence with the laser-Doppler anemometer," *Annu. Rev. Fluid Mech.* **11**, 443–503 (1979).
- ²⁸W. L. Nyborg, "Acoustic streaming," in *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, San Diego, 1998), Chap. 7, pp. 207–231.
- ²⁹This method was used by Biwa *et al.* (Ref. 20) in their study of acoustic-streaming velocity generated by planar traveling waves in a cylindrical waveguide. In such a traveling wave, the physical difference between the Eulerian and Lagrangian streaming velocities is substantial. This difference is in addition to any instrumentation error that is present due to the LDA/BSA system. Even so, the Lagrangian streaming velocity will still be accurately determined using the method developed in the present study.
- ³⁰The measurements of Eulerian streaming velocity shown in Fig. 8 were reported previously by Thompson and Atchley (Refs. 6, 7). In the present study, the sampled velocity data are reanalyzed in order to determine the Lagrangian streaming velocities.

Influences of a temperature gradient and fluid inertia on acoustic streaming in a standing wave^{a)}

Michael W. Thompson,^{b)} Anthony A. Atchley,^{c)} and Michael J. Maccarone

Graduate Program in Acoustics, The Pennsylvania State University, University Park, Pennsylvania 16802

(Received 1 August 2004; revised 27 November 2004; accepted 29 December 2004)

Following the experimental method of Thompson and Atchley [J. Acoust. Soc. Am. **117**, 1828–1838 (2005)] laser Doppler anemometry (LDA) is used to investigate the influences of a thermoacoustically induced axial temperature gradient and of fluid inertia on the acoustic streaming generated in a cylindrical standing-wave resonator filled with air driven sinusoidally at a frequency of 308 Hz. The axial component of Lagrangian streaming velocity is measured along the resonator axis and across the diameter at acoustic-velocity amplitudes of 2.7, 4.3, 6.1, and 8.6 m/s at the velocity antinodes. The magnitude of the axial temperature gradient along the resonator wall is varied between approximately 0 and 8 K/m by repeating measurements with the resonator either surrounded by a water jacket, suspended within an air-filled tank, or wrapped in foam insulation. A significant correlation is observed between the temperature gradient and the behavior of the streaming: as the magnitude of the temperature gradient increases, the magnitude of the streaming decreases and the shape of the streaming cell becomes increasingly distorted. The observed steady-state streaming velocities are not in agreement with any available theory. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1859992]

PACS numbers: 43.25.Nm, 43.25.Zx [MFH]

Pages: 1839–1849

I. INTRODUCTION

The acoustic streaming generated by a monofrequency acoustic standing wave in a long cylindrical resonator has been studied experimentally by Thompson and Atchley¹ using the laser Doppler anemometry (LDA) technique. At relatively low acoustic amplitudes, the measured axial component of the outer streaming-velocity field is in agreement with the theory of Rott:² the variation with respect to the axial coordinate is sinusoidal, the variation with respect to the transverse coordinate is parabolic, and the variation with respect to the acoustic amplitude is quadratic. These results are consistent with previous measurements made in the vicinity of the acoustic-velocity node by Sharpe *et al.*,³ Arroyo and Greated,⁴ Hann and Greated,⁵ and Rockliff *et al.*⁶ using the particle image velocimetry (PIV) technique.

Thompson and Atchley⁷ previously observed that the variation of the axial component of streaming velocity with respect to the transverse coordinate is not parabolic at relatively large acoustic amplitudes. They noted the qualitative consistency between their measurements and the computational results of Menguy and Gilbert,⁸ who studied the influence of fluid inertia on acoustic streaming.

In thermoacoustic engines and refrigerators, the pres-

ence of any acoustic streaming within the acoustical resonator tends to reduce the efficiency of the device.⁹ The influence of an axial temperature gradient on acoustic streaming has been considered theoretically by Rott,² Olson and Swift,¹⁰ and Bailliet *et al.*¹¹

In the present study, the influences of a thermoacoustically induced axial temperature gradient and of fluid inertia on the outer streaming-velocity field generated by a monofrequency acoustic standing wave in a long cylindrical resonator are investigated experimentally using the method of Thompson and Atchley:¹ the fluid velocity is measured using laser Doppler anemometry with burst spectrum analysis (BSA), and the Lagrangian streaming velocity is determined using the observed acoustic motion and the arrival times of the velocity samples.

The relevant theories are reviewed in Sec. II, the experimental procedure is described in Sec. III, and measurements are presented in Sec. IV.

II. THEORY

A. Influence of a temperature gradient on streaming

Rott² derived an expression for the axial component of the streaming-velocity field generated in a long cylindrical tube filled with a fluid supporting a monofrequency acoustic standing wave. The coordinate system necessary to describe the fluid motion consists of an axial coordinate running parallel to the tube centerline and a transverse coordinate running perpendicular to the centerline; the fluid motion is assumed to be symmetric about the centerline. The axial component of acoustic velocity outside the acoustic boundary layer has the form

^{a)}Portions of this work were presented in: M. W. Thompson and A. A. Atchley, "Optical measurement of acoustic streaming in a standing wave with a temperature gradient," 147th Meeting of the Acoustical Society of America, NY, May 2004; and M. W. Thompson, "Measurement of acoustic streaming in a standing wave using laser Doppler anemometry" (Ph.D. dissertation, The Pennsylvania State University, University Park, PA, 2004).

^{b)}Current affiliation: Applied Physics Laboratory, The Johns Hopkins University, Laurel, Maryland 20723.

Electronic mail: michael.wayne.thompson@jhuapl.edu

^{c)}Electronic mail: atchley@psu.edu

$$u_1 = \Re(\tilde{U}_1 e^{j\omega t}) \sin\left(\frac{2\pi x}{\lambda}\right), \quad (1)$$

where \tilde{U}_1 is the complex amplitude of the acoustic velocity at its antinode, ω is the angular acoustic frequency, t is time, x is the axial coordinate, λ is the acoustic wavelength, $j = \sqrt{-1}$, and \Re denotes taking the real part of the complex argument. The transverse component of acoustic velocity is negligible outside the acoustic boundary layer.

In his derivation, Rott assumed: (1) that the inner radius of the resonator, R , is small relative to the acoustic wavelength λ but large relative to the thicknesses of both the viscous acoustic boundary layer, $\delta_\nu = \sqrt{2\nu/\omega}$, and the thermal acoustic boundary layer, $\delta_\kappa = \sqrt{2\kappa/\omega}$, where ν and κ are the kinematic viscosity and thermal diffusivity of the fluid; (2) that terms higher than second order in the acoustic Mach number $M = U_1/c$ are negligible, where U_1 is the magnitude of \tilde{U}_1 and c is the sound speed in the fluid; (3) a relationship between temperature T and viscosity μ of the form $\mu \propto T^\beta$, where β is a constant that depends on the properties of the fluid; and (4) a variation of mean temperature T_0 with respect to the axial coordinate x .

Rott's expression for the axial component of streaming velocity outside the acoustic boundary layer is

$$\begin{aligned} \langle u_2 \rangle = & (1 + \alpha_1) \frac{3}{8} \frac{U_1^2}{c} \left(1 - \frac{2r^2}{R^2}\right) \sin\left(\frac{\pi x}{L}\right) \\ & - \alpha_2 \frac{\theta u_1^2}{2\omega} \left(1 - \frac{2r^2}{R^2}\right), \end{aligned} \quad (2)$$

where r is the transverse coordinate ($-R < r < R$); $L = \lambda/4$ is the length of a streaming cell in the axial direction; u_1 is the axial component of acoustic velocity given by Eq. (1); and α_1 , α_2 , and θ are given by

$$\alpha_1 = \frac{2}{3} (1 - \beta)(\gamma - 1) \frac{\sqrt{\text{Pr}}}{1 + \text{Pr}}, \quad (3a)$$

$$\alpha_2 = \frac{(1 - \beta)(1 - \sqrt{\text{Pr}})}{2(1 + \sqrt{\text{Pr}})(1 + \text{Pr})}, \quad (3b)$$

$$\theta = \frac{1}{T_0} \frac{dT_0}{dx}, \quad (3c)$$

where γ is the ratio of the isobaric to isochoric specific heats and $\text{Pr} = \nu/\kappa$ is the Prandtl number. (This same result, for the case $\beta=0$ and $\theta=0 \text{ m}^{-1}$, was later obtained independently by Qi.¹²) Typical values of α_1 and α_2 in air ($\gamma=1.4$, $\text{Pr}=0.71$, $\beta=0.77$) are $\alpha_1=0.030$ and $\alpha_2=0.0057$.

For the conditions corresponding to the largest temperature gradient used in the present study ($dT_0/dx \approx 8 \text{ K/m}$, $T_0 \approx 297 \text{ K}$, $\theta \approx 0.03 \text{ m}^{-1}$, $U_1 = 8.5 \text{ m/s}$, $c \approx 346 \text{ m/s}$, $\omega = 1940 \text{ s}^{-1}$), the magnitude of the first term on the right-hand side of Eq. (2) has a maximum value of 8.1 cm/s , whereas the magnitude of the second term has a maximum value that is four orders of magnitude smaller: $3 \text{ } \mu\text{m/s}$. The second term is therefore negligible relative to the first and the axial streaming velocity given by Eq. (2) can be rewritten as

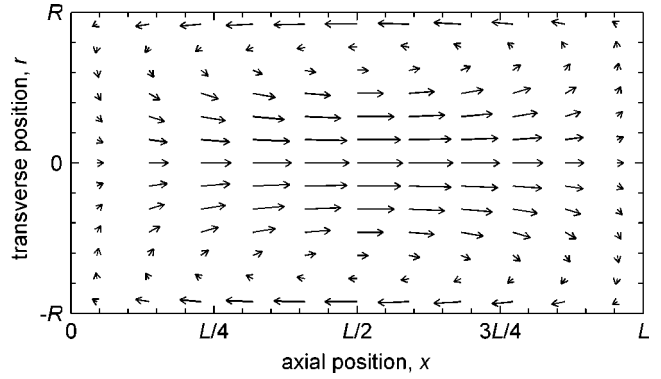


FIG. 1. Profile view of a streaming cell in the outer streaming-velocity field for the case of a relatively small axial temperature gradient. [The axial component of streaming velocity is calculated from Eq. (4).] The flow along the resonator walls ($r = \pm R$) is toward the acoustic-velocity node ($x = 0$), and the return flow along the centerline ($r = 0$) is toward the acoustic-velocity antinode ($x = L$). Neighboring cells, to the left and to the right of the cell shown here, circulate in the opposite direction. As r approaches $\pm R$ within the boundary layer (not shown), the velocity goes to zero. (Reprinted from Ref. 1 with permission.)

$$\langle u_2 \rangle = (1 + \alpha_1) \frac{3}{8} \frac{U_1^2}{c} \left(1 - \frac{2r^2}{R^2}\right) \sin\left(\frac{\pi x}{L}\right). \quad (4)$$

In this simplified expression, the variation of axial streaming velocity with respect to the axial coordinate x is sinusoidal, with the maximum value occurring at $x = L/2$; the variation with respect to the transverse coordinate r is parabolic, with the average flow over the cross section (x held constant) equal to zero; and the variation with respect to the acoustic amplitude U_1 is quadratic. Figure 1 shows a profile view of the theoretical streaming-velocity field within a single streaming cell (including the transverse component, which is not given here).

Olson and Swift¹⁰ extended Rott's theory² by deriving an expression for the axial component of the time-averaged mass-flux density generated by an acoustic field that is an arbitrary combination of monofrequency standing and traveling waves in a long resonator having a circular cross-sectional area that varies with respect to the axial coordinate x . For the case of a pure standing wave in a cylindrical resonator, their results are equivalent to those of Rott.

Bailliet *et al.*¹¹ derived equations from which the axial component of streaming velocity can be calculated numerically for arbitrary values of the ratio R/δ_κ . For the case of a monofrequency standing wave in a cylindrical resonator with $R \gg \delta_\kappa$ and with an axial temperature gradient on the order of $\theta = (dT_0/dx)/T_0 \sim 1 \text{ m}^{-1}$, they showed that the expected magnitude of the axial streaming velocity along the centerline ($r = 0$) is in agreement with the value predicted by Olson and Swift,¹⁰ while the variation of the axial streaming velocity with respect to the transverse coordinate r appears to be approximately parabolic. Although no results were given for the case of a relatively small temperature gradient (such as those used in the present study: $\theta \lesssim 0.03 \text{ m}^{-1}$), such results would also be expected to be in agreement with the theory of Olson and Swift.

B. Influence of fluid inertia on streaming

Lighthill¹³ discussed the influence of fluid inertia on acoustic streaming, concluding that the effect is significant in nearly all noticeable acoustic streaming motions. As the inertia of a fluid parcel being convected about the streaming cell becomes more significant at larger streaming velocities, it becomes more difficult for the magnitude and direction of the parcel's velocity to be changed.

The influence of fluid inertia on the shape of the streaming-velocity field generated by a monofrequency acoustic standing wave in a cylindrical resonator was studied computationally by Menguy and Gilbert.⁸ Following the work of Qi,¹² they neglected the variation of viscosity and thermal conductivity with respect to temperature ($\beta=0$) and assumed a uniform axial temperature distribution ($\theta=0$ m⁻¹).

They related the axial and transverse components of streaming velocity by considering the time-averaged equation of mass conservation to second order in the acoustic Mach number M and the time-averaged equation of momentum conservation to fourth order in M . In the equation of momentum conservation, the relationship between the axial and transverse components of streaming velocity is shown to scale according to the Reynolds number

$$\text{Re} = \left(\frac{U_1}{c} \right)^2 \left(\frac{R}{\delta_\nu / \sqrt{2}} \right)^2. \quad (5)$$

The case $\text{Re} \ll 1$ corresponds to the relatively slow streaming considered by Rott,² Qi,¹² Olson and Swift,¹⁰ and Bailliet *et al.*,¹¹ in which the variation of axial streaming velocity with respect to the transverse coordinate r is found to be parabolic using only the time-averaged equation of mass conservation.

The case $\text{Re} \geq 1$ is referred to as “nonlinear streaming.” Menguy and Gilbert's method for calculating the nonlinear outer streaming-velocity field proceeds as follows: The influence of fluid inertia within the boundary layer is first shown to be negligible. The axial component of streaming velocity calculated at the edge of the boundary layer is then used as a boundary condition in the calculation of the nonlinear outer streaming-velocity field. The axial and transverse components of the outer field are each expressed as the product of a Fourier series involving the axial coordinate x and a power series involving the transverse coordinate r , and the conservation equations are solved for the coefficients in the resulting expressions.

The results of these calculations predict that the fluid inertia at moderate Reynolds numbers ($1 \leq \text{Re} \leq 4$) distorts the shape of the outer streaming cells by causing the location of maximum streaming velocity along the centerline ($r=0$) to shift from the midpoint ($x=L/2$) toward the velocity antinode ($x=L$). As a result of this distortion, the variation of the axial component of streaming velocity with respect to the axial coordinate x is no longer sinusoidal and the variation with respect to the transverse coordinate r is no longer parabolic. The Reynolds number Re determines the degree to which the streaming-velocity field is distorted relative to the field in the slow-streaming case.

C. Thermoacoustic heat transport

For the case considered by Rott²—a monofrequency acoustic standing wave in a long cylindrical resonator—heat is transported thermoacoustically along the inner wall of the resonator from the acoustic-velocity antinodes toward the acoustic-velocity nodes. Merkli and Thomann¹⁴ worked out a basic theory predicting the time-averaged heat flux between the resonator wall and the enclosed fluid and made measurements confirming that the wall in the vicinity of the antinodes is cooled, while the wall in the vicinity of the nodes is heated.

A more recent treatment of this phenomenon has been given by Swift.¹⁵ An admittedly oversimplified explanation is as follows: Consider a parcel of fluid oscillating parallel to the resonator centerline at a distance of approximately one thermal penetration depth δ_κ from the resonator wall, which initially has a uniform temperature distribution, at a location somewhere between an acoustic-velocity antinode ($x=L$) and an acoustic-velocity node ($x=0$). At the extreme of the parcel's motion that is nearest the antinode, the parcel's volume is large and its temperature is lower than the wall temperature, causing net heat to flow from the wall into the parcel. The parcel transports this increase in entropy toward the other extreme of its motion that is nearest the node. At this location, the parcel's volume is small and its temperature is higher than the wall temperature, causing net heat to flow from the parcel into the wall. The parcel returns to its initial position and the process is repeated.

Due to the action of similar fluid parcels along the entire length of the resonator, a sort of “bucket brigade” is formed that transports net heat along the inner wall from the velocity antinodes toward the velocity nodes, causing the wall temperatures at the antinodes to decrease and the wall temperatures at the nodes to increase. Under most circumstances, the rate of thermoacoustic heat transport is eventually balanced by the rate at which heat is transported by other means (such as conduction axially within the resonator wall or conduction and convection radially outward from the resonator into the surrounding environment), and a steady-state axial temperature gradient is reached, the calculation of which is non-trivial. The thermoacoustic heat transport and the resultant steady-state temperature gradient are both proportional to U_1^2 , the square of the acoustic-velocity amplitude.

III. PROCEDURE

A. Experimental apparatus

The apparatus shown in Fig. 2 has been used to study the influences of a thermoacoustically induced temperature gradient and of fluid inertia on the outer streaming-velocity field generated by a monofrequency acoustic standing wave in air. It is identical to the apparatus used previously by Thompson and Atchley;¹ the details of its construction will not be repeated here.

The resonator is filled with air ($T_0 \approx 23^\circ\text{C}$, $c \approx 345$ m/s) in which a small amount of smoke from burned cotton rope (particle diameter ≈ 1 μm) has been introduced. A standing wave is produced by driving the two compression drivers (University Sound 1829BT) in phase at the fourth

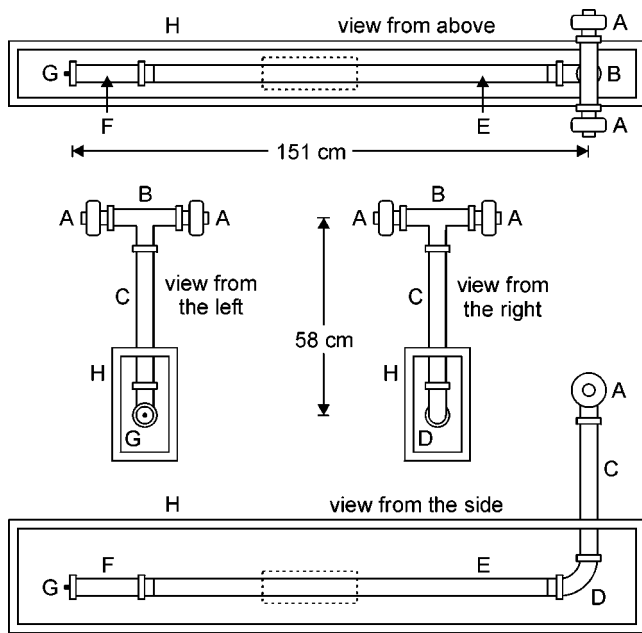


FIG. 2. The experimental apparatus. (A) Two compression drivers are connected to (B) a stainless-steel tee on top of (C) a 40 cm stainless-steel tube. This vertical section is connected by (D) a stainless-steel 90 deg elbow to a horizontal section comprised of (E) a 122 cm borosilicate-glass tube and (F) a 20 cm stainless-steel tube. The resonator is terminated by (G) a rigid stainless-steel end cap. The horizontal section is suspended inside (H) a glass-walled enclosure. The dotted box indicates the location of the streaming cell where measurements are made, which is similar to the cell shown in Fig. 1. (Reprinted from Ref. 1 with permission.)

natural frequency of the air column ($f = \omega/2\pi = 308$ Hz, $\lambda = 1.13$ m). The smoke follows the motion of the air, serving as tracer particles for the velocity measurements. The relative humidity inside the resonator is not significantly affected by the presence of the smoke [verified using a portable thermohygrometer (TH Pen 8708)].

The axial component of fluid velocity within the long glass section of the resonator (inner diameter $2R = 4.65$ cm, wall thickness = 2.0 mm) is sampled as a function of time using a single-component laser Doppler anemometry (LDA) system (Dantec Measurement Technology FlowLite) connected to a burst spectrum analysis (BSA) system (Dantec Measurement Technology 57N11). The absolute accuracy of the velocity measurements is approximately 0.3%, based on the uncertainty in the intersection angle of the two LDA laser beams. However, the random errors in the individual velocity samples are much larger than 0.3% and vary with the BSA settings used. The LDA probe is mounted on a linear positioning system that translates in both the axial and the transverse directions in the horizontal plane.

For the largest acoustic amplitude used in the present study ($U_1 = 8.6$ m/s, $Re = 40$), the amplitude of the second harmonic of velocity at its antinode is approximately 10% of the amplitude of the first harmonic at its antinode. Even so, the assumption of a monofrequency acoustic field is still expected to be valid because the magnitude of the streaming generated by this second-harmonic standing wave is expected to be only 1% of the magnitude of the streaming generated by the first harmonic, and therefore too small to be detected.

The amplitude of the first harmonic of pressure at its antinode is monitored using a piezoresistive gauge-pressure transducer (Endevco 8510B-5) mounted flush in the rigid end cap. A barometric-pressure transducer (SenSymb ASCX15AN) is located within the glass-walled enclosure so that, in combination with the gauge-pressure transducer, the mean absolute pressure within the resonator (nominally 97 kPa) can be determined.

Several type-E thermocouples (Omega Engineering TT-E-30-SLE) are attached to the outer wall of the resonator at various locations in order to monitor the spatial distribution of mean temperature. Each thermocouple consists of a twisted wire pair that is soldered together and attached to the resonator using a $1.3 \text{ cm} \times 1.3 \text{ cm}$ piece of copper shielding tape that is then covered with transparent tape. The accuracy of the temperature measurements is approximately 0.1°C , determined using an ice-point bath at 0.0°C .

In order to vary the magnitude of the thermoacoustically induced axial temperature gradient while holding the Reynolds number Re constant, three different boundary conditions are imposed on the wall of the resonator.

- (1) In order to produce an insulated boundary condition, the glass section of the resonator is wrapped in two layers of 1.0-cm-thick polyethylene pipe insulation. The insulation extends a distance of approximately 90 cm in the axial direction and is centered on the streaming cell of interest. The laser beams of the LDA system pass into the resonator through a 0.8-cm-high slit in the insulation (at $r = R$) that extends 40 cm in the axial direction and is centered on the streaming cell of interest.
- (2) In order to produce an isothermal boundary condition, the glass section is replaced with an identical tube mounted inside a water jacket. The jacket consists of a larger-diameter borosilicate-glass tube mounted concentric to the smaller-diameter resonator tube using custom-built PVC collars at each end. The outer tube has an inner diameter of 7.5 cm and a wall thickness of 2.5 mm, resulting in a radial spacing of 1.2 cm between the outer wall of the inner tube and the inner wall of the outer tube. Water enters the jacket through an orifice in the collar at the end farthest from the drivers and exits through an orifice in the collar at the end nearest the drivers. Flexible reinforced PVC tubing is used to connect the jacket to a circulator bath (VWR Scientific 1156) that circulates water through the jacket at a rate of 7 L/min while maintaining the water at a constant temperature.
- (3) The third boundary condition corresponds to the glass section of the resonator being surrounded simply by air, and will be referred to as the “uncontrolled” boundary condition. Although this case is more difficult to model mathematically, it serves the purpose of producing an intermediate axial temperature gradient with a magnitude that lies between the insulated and isothermal gradients for a given Reynolds number Re .

Between sets of streaming measurements, the smoke residue is cleaned from the inner wall of the resonator using an ammonia-based household glass cleaner. In preliminary

experiments done with an isothermal boundary condition, spurious results were observed when measurements were made shortly after the resonator had been cleaned. To avoid this problem, the resonator is allowed to dry for several hours before measurements are made.

B. Determination of streaming velocity

The sample records of fluid velocity obtained using the LDA/BSA technique are postprocessed according to the method of Thompson and Atchley.¹ For each record, a complex Fourier series representing the velocity waveform is first calculated using the Fourier averaging (FA) method developed by Sonnenberger *et al.*¹⁶ The resultant Fourier coefficients corresponding to the time-harmonic waveform components are then used in conjunction with the arrival times of the velocity samples to determine the Lagrangian streaming velocities of individual tracer particles by imposing the constraint that the location where the particle velocities are sampled (the LDA fringe volume) is spatially fixed.¹ If the streaming changes negligibly with respect to time during a sample record, then a single value for the Lagrangian streaming velocity is determined that best represents all of the tracer particles within the record. The uncertainties in the Lagrangian streaming velocities determined using this method depend primarily on the random errors in the individual velocity samples. This method avoids an offset error observed in the Eulerian streaming velocity determined using the FA method. The source of this error is attributed to the LDA/BSA system.

The above-outlined method assumes that the Eulerian and Lagrangian acoustic-velocity fields are equivalent. For an acoustic standing wave of the type considered in the present study, the difference between the Eulerian and Lagrangian velocities amounts to a correction to the magnitude and phase of the second harmonic of velocity.¹ The magnitude of this correction is at most $U_1^2/4c$, which, for the largest acoustic amplitude used in the present study ($U_1 = 8.6$ m/s, $Re=40$), is only approximately 0.6% of the magnitude of the first harmonic. Such a small correction would produce a negligible change in the magnitude and shape of the velocity waveform (negligible for the purpose of determining the Lagrangian streaming velocity), and the correction can therefore be safely ignored.

IV. MEASUREMENTS

A. Influence of a temperature gradient

Figure 3 shows the observed correlation between the thermoacoustically induced axial temperature gradient along the resonator wall and the axial component of streaming velocity for the three boundary conditions: insulated, isothermal, and uncontrolled. When the acoustic field is switched on at time $t=0$ min, the axial temperature distribution is initially uniform [see Fig. 3(b)]. Due to thermoacoustic heat transport along the inner wall of the resonator, the wall temperature at the acoustic-velocity node ($x=0$) increases, as a function of time, relative to the temperature at the antinode ($x=L$). As the magnitude of the temperature gradient increases, the magnitude of the axial component of streaming

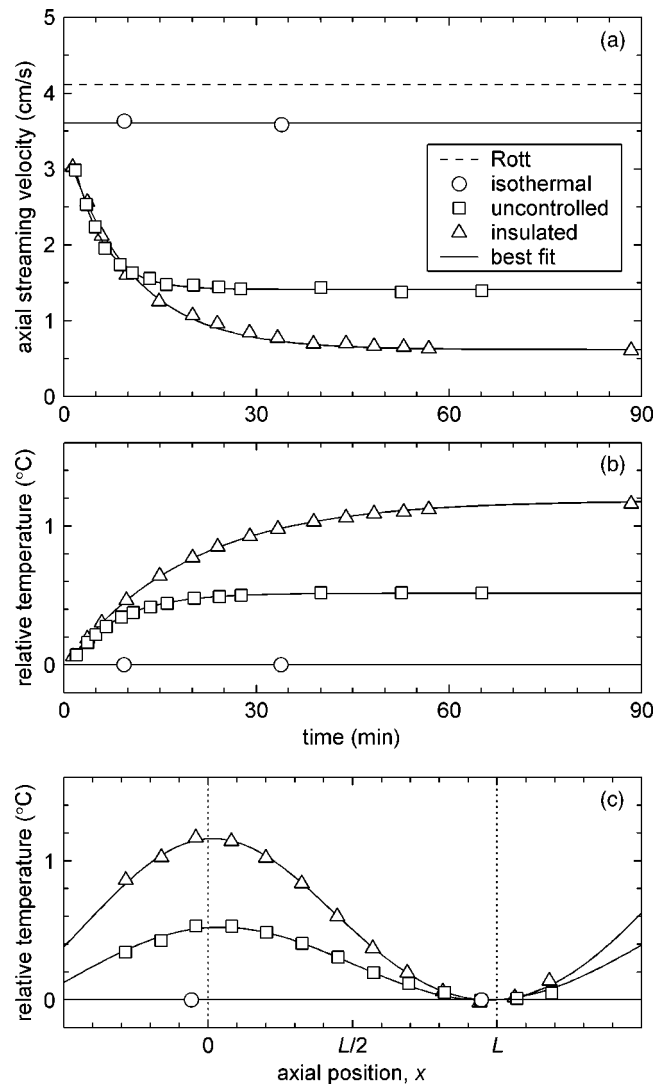


FIG. 3. In (a), the axial components of streaming velocity ($U_1=6.1$ m/s, $Re=20$) for the three boundary conditions, measured as functions of time at the center of the streaming cell ($x=L/2$, $r=0$), are compared to the theoretical streaming velocity predicted by Rott [see Eq. (4)]. In (b), corresponding measurements of the external wall temperatures at the acoustic-velocity node ($x=0$) relative to the temperatures at the antinode ($x=L$) are compared. As the magnitudes of the temperature gradients in (b) increase, the corresponding magnitudes of the streaming velocities in (a) decrease. In (c), the steady-state axial temperature distributions relative to the temperatures at the velocity antinode ($x=L$) are compared.

velocity decreases [see Fig. 3(a)] until a steady-state condition has eventually been reached. (The value of streaming velocity shown here at time $t=0$ min corresponds to the “steady-state” value of streaming velocity discussed by Thompson and Atchley.¹ This value is reached approximately 5 s after the acoustic field is switched on.)

For the case of an uncontrolled boundary condition, the axial streaming velocity reaches 95% of its steady-state value within approximately 14 min and the axial temperature gradient reaches 95% of its steady-state condition within approximately 23 min. For the case of an insulated boundary condition, the times are approximately 32 min for the axial streaming velocity and approximately 54 min for the axial temperature gradient. These times are independent of the acoustic amplitude. Additionally, the results of this experi-

ment do not depend significantly on whether the resonator is completely sealed with respect to the laboratory or whether a small leak is present.

The temperature-dependent behavior of the streaming observed in Fig. 3 is puzzling because it is not predicted by any of the available theories discussed in Sec. II. The largest discrepancies between the measured streaming velocities and the theory of Rott² are observed when using an insulated boundary condition, which produces the largest steady-state temperature gradient.

This experiment has been repeated with an initially large temperature gradient that is produced by switching on the acoustic field, allowing the temperature gradient and the streaming to reach steady-state conditions, and then switching off the acoustic field so that the fluid comes completely to rest but the temperature gradient remains. The acoustic field is then switched on again, and the streaming velocity observed just a few seconds later is in agreement with the steady-state value it had immediately before the acoustic field was switched off. This indicates that it is the temperature gradient that controls the streaming, and not *vice versa*.

The temperatures reported for the insulated and uncontrolled boundary conditions in Figs. 3(b) and (c) are determined using thermocouples attached along the top outer surface of the glass resonator tube. Corresponding thermocouples attached along the bottom outer surface give temperature readings that are lower than the top temperatures by less than 0.1 °C, indicating that there is occasionally a slight vertical temperature gradient within the resonator.

In a previous experiment done with an uncontrolled boundary condition, thermocouples were attached to the inner and outer surfaces of the glass tube near an acoustic-velocity node (where the temperature difference between the resonator wall and the surrounding air is greatest). For the largest acoustic amplitude used in the present study ($U_1 = 8.6$ m/s, $Re=40$), the inner temperature is less than 0.1 °C higher than the outer temperature, while the outer temperature difference between the two ends of the streaming cell is 2.3 °C. The temperature difference between the inner and outer surfaces of the glass tube near an acoustic-velocity antinode is negligible.

For the case of an isothermal boundary condition, thermocouples were originally attached to the outer surface of the resonator inside the water jacket. For the largest acoustic amplitude used in the present study ($U_1 = 8.6$ m/s, $Re=40$), the variation of steady-state temperature with respect to the axial coordinate x was observed to be uniform within 0.1 °C. However, these thermocouples eventually began giving spurious temperature readings because air was being drawn through the insulation surrounding the thermocouple wires and into the water jacket. So, the wires were cut and sealed to prevent further problems. [The observations of temperature reported for the isothermal boundary condition in Figs. 3(b) and (c) are set equal to zero and are included solely for illustrative purposes.]

For the case of an insulated boundary condition, a noticeable increase in local temperature was observed if the LDA laser beams were permitted to shine for prolonged periods of time on the inner surface of the insulation covering

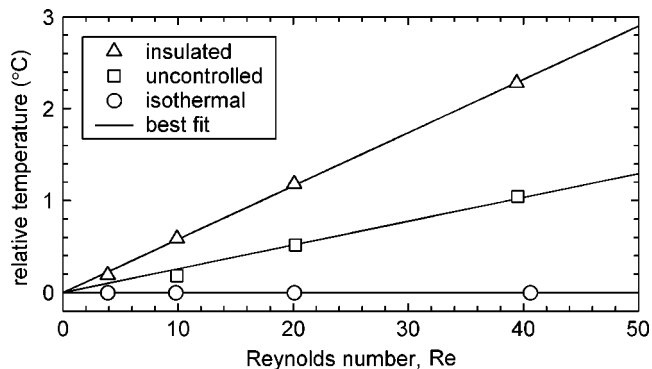


FIG. 4. Steady-state temperatures at the acoustic-velocity node ($x=0$) relative to the temperatures at the antinode ($x=L$) as a function of Reynolds number Re and boundary condition.

the far side of the resonator (the side at $r = -R$, opposite the slit and the LDA probe). Therefore, a shutter mechanism is used to temporarily block the beams whenever a measurement of velocity is not being made.

Figure 4 shows the steady-state axial temperature difference across the streaming cell as a function of Reynolds number Re and boundary condition. For each boundary condition, the data show good agreement with a linear fit. This is expected from the theory (see Sec. II) because the thermoacoustic heat transport, the steady-state axial temperature gradient, and the Reynolds number are each proportional to U_1^2 , the square of the acoustic-velocity amplitude. The value reported at $Re=10$ for the uncontrolled boundary condition is artificially low by about 0.1 °C due to an opposing linear temperature gradient extending axially along the resonator. (The temperature differences reported for the isothermal boundary condition are set equal to zero and are included solely for illustrative purposes.)

B. Steady-state streaming

Figure 5 shows the axial component of the steady-state streaming-velocity field at a Reynolds number of $Re=4$ for isothermal and insulated boundary conditions. In Fig. 5(a), there is a noticeable difference between the isothermal and insulated data even though the steady-state axial temperature difference across the streaming cell is only 0.2 °C for the insulated boundary condition (see Fig. 4). This behavior is repeatable and is not predicted by any of the available theories discussed in Sec. II.

Previous measurements made by Thompson and Atchley¹ at a Reynolds number of $Re=0.36$ showed that the magnitude of the streaming-velocity field is noticeably influenced by the dependence of viscosity and thermal conductivity on temperature, consistent with the theory of Rott² (see Sec. II A). However, Menguy and Gilbert⁸ neglected this effect (see Sec. II B), causing their computationally predicted streaming velocities to be larger than Rott's predictions by a factor of 1.097. In order to correct for this discrepancy, the results of Menguy and Gilbert shown in Fig. 5 have been scaled by a factor of $1/1.097=0.912$. Even so, the isothermal data shown in Fig. 5 are in better agreement with the theory

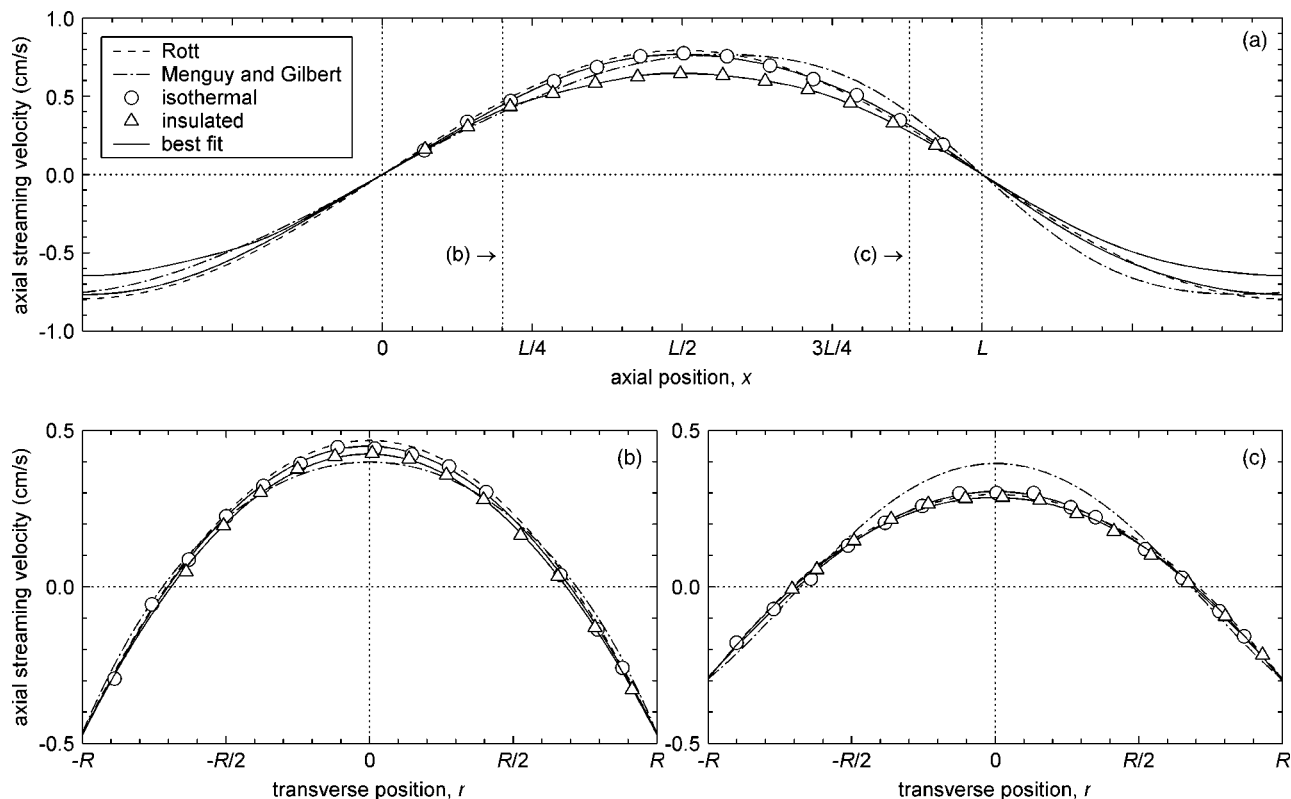


FIG. 5. Axial component of the steady-state streaming-velocity field at a Reynolds number of $Re=4$ ($U_1=2.66$ m/s) for isothermal and insulated boundary conditions. In (a), the variation of axial streaming velocity with respect to the axial coordinate x ($r=0$) is compared to the theories of Rott [see Eq. (4)] and of Menguy and Gilbert (Ref. 8) (see Sec. II B). In (b) and (c), the variation of axial streaming velocity with respect to the transverse coordinate r ($x=0.20L$ and $0.88L$) is compared to theory.

of Rott, suggesting that the influence of fluid inertia on the streaming-velocity field is not as significant as predicted by Menguy and Gilbert.

Figure 6 shows the axial component of the steady-state streaming-velocity field at a Reynolds number of $Re=10$ for all three boundary conditions: insulated, isothermal, and uncontrolled. The influence of the axial temperature gradient on the streaming is noticeable: when the steady-state axial temperature difference across the streaming cell is increased by using an uncontrolled or insulated boundary condition (see Fig. 4), the discrepancies between the measured streaming velocities and the theory of Rott also increase. Again, this behavior is repeatable and is not predicated by any of the available theories discussed in Sec. II.

For the isothermal data shown in Fig. 6(a), the location of maximum streaming velocity along the centerline is shifted slightly toward the acoustic-velocity antinode ($x=L$). This behavior is qualitatively consistent with the influence of fluid inertia. However, Menguy and Gilbert⁸ do not give any results for $Re>4$ because of convergence problems with their computational method. Additionally, their analysis is not expected to be valid for $Re\geq 10$.

Figures 7 and 8 show the axial components of the steady-state streaming-velocity fields at Reynolds numbers of $Re=20$ and 40 . As the Reynolds number and the axial temperature gradient are increased, the discrepancies between the measured data and the theory become even more pronounced. [The uncontrolled data in Fig. 7(c) are in agreement with the measurements reported previously for this lo-

cation and approximate Reynolds number ($Re=19.3$) by Thompson and Atchley.⁷]

For the insulated data in Fig. 7, the magnitude of the streaming velocity near the position $x=0.6L$ and $r=0$ appears to be a local minimum. This trend is continued in Fig. 8, where the streaming along a large section of the centerline ($r=0$) has been completely suppressed for both insulated and uncontrolled boundary conditions. At $Re=40$, the streaming velocity at the center of the cell ($x=L/2$, $r=0$), measured as a function of time, goes to zero within approximately 9 min after the acoustic field is switched on.

The insulated data in Figs. 7 and 8 exhibit some asymmetry with respect to the transverse coordinate r . This is most likely due to slight differences in temperature between the resonator walls at $r=\pm R$ caused by the slit in the insulation on the wall at $r=R$ (see Sec. III A). For the isothermal data in the vicinity of the acoustic-velocity antinode ($x=L$) in Figs. 7(a) and 8(a), the curves fit to the measured data (using a Fourier sine series) are troublesome and have therefore been omitted.

For each combination of Reynolds number and boundary condition considered in Figs. 5–8, the extent to which the measurements are consistent with the principle of conservation of mass can be determined by averaging the transverse profile of axial streaming velocity over the resonator cross section at each axial position. For the data shown in Figs. 5(b)–(c), 6(b)–(d), 7(b)–(d), and 8(b)–(d), the computed cross-sectional average velocities are on the order of only 1% of the magnitudes of the streaming velocities near

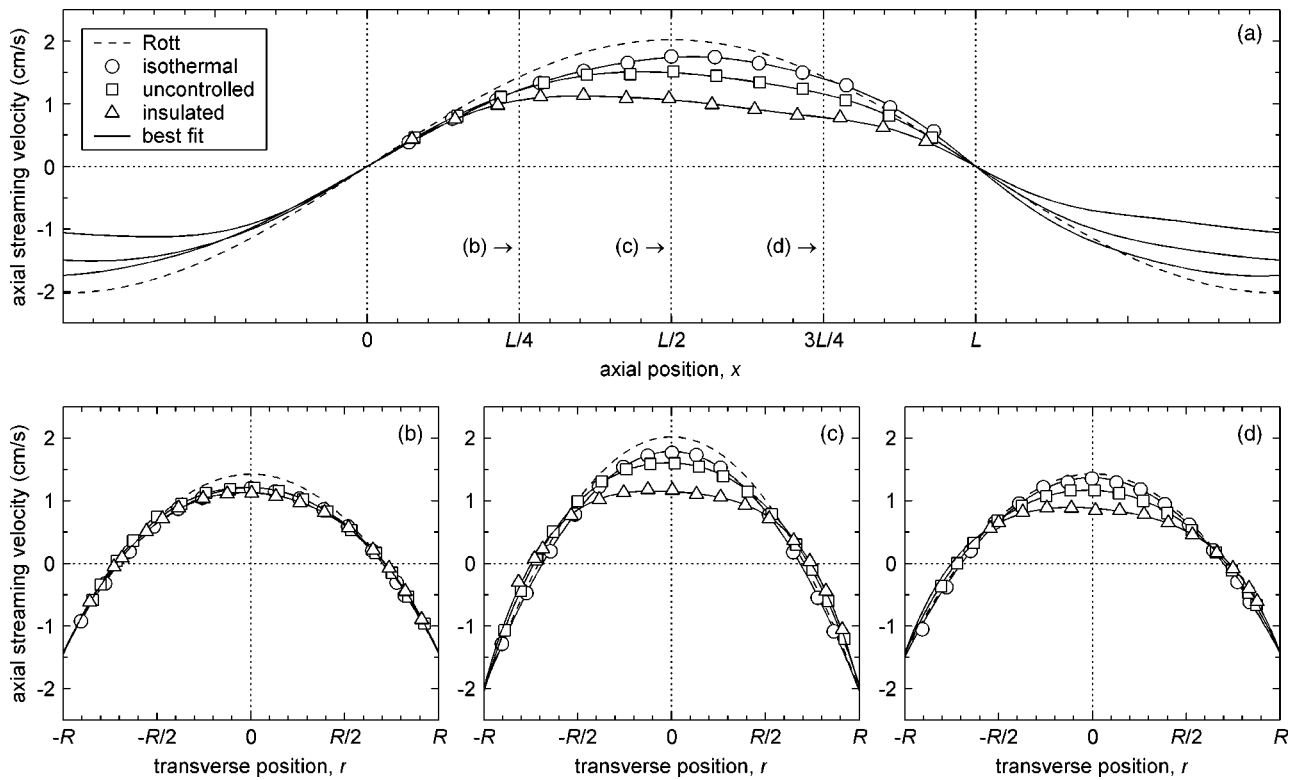


FIG. 6. Axial component of the steady-state streaming-velocity field at a Reynolds number of $Re=10$ ($U_1=4.26$ m/s) for all three boundary conditions. In (a), the variation of axial streaming velocity with respect to the axial coordinate x ($r=0$) is compared to the theory of Rott [see Eq. (4)]. In (b), (c), and (d), the variation of axial streaming velocity with respect to the transverse coordinate r ($x=L/4$, $L/2$, and $3L/4$) is compared to theory.

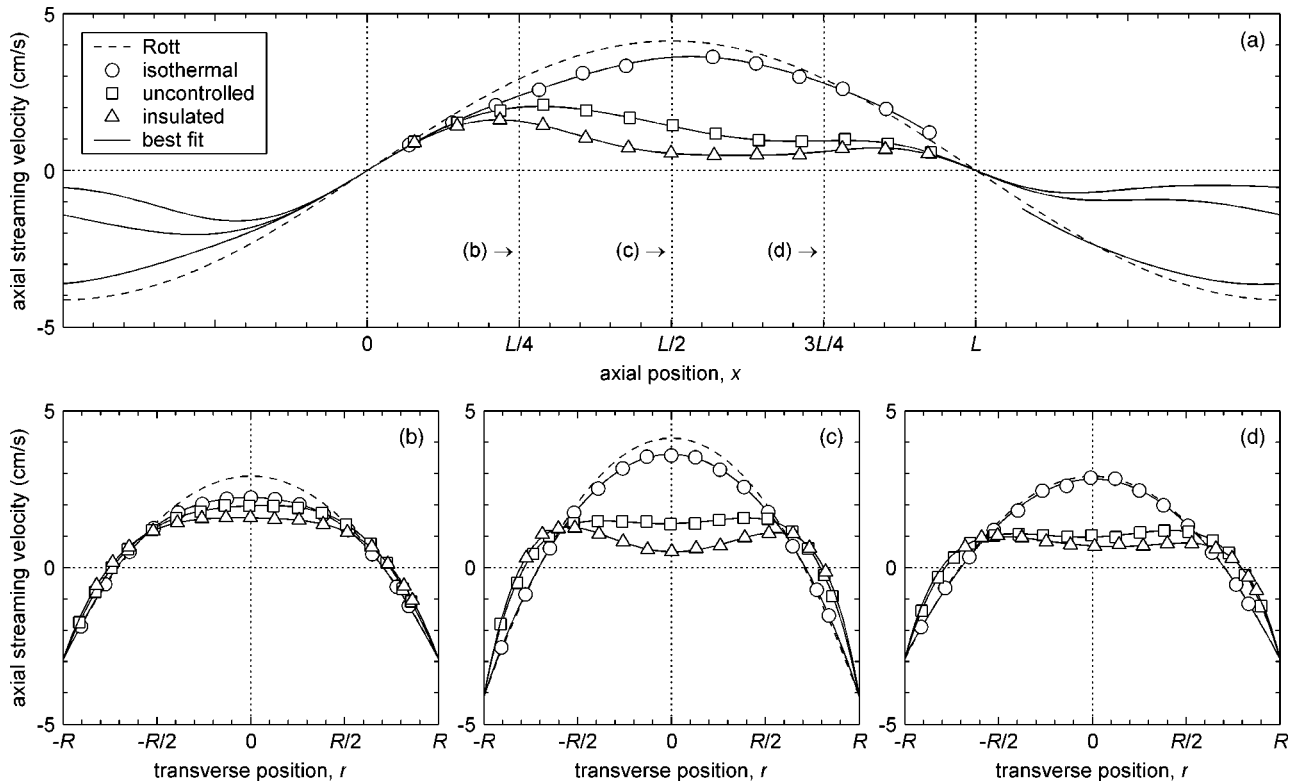


FIG. 7. Axial component of the steady-state streaming-velocity field at a Reynolds number of $Re=20$ ($U_1=6.11$ m/s) for all three boundary conditions. In (a), the variation of axial streaming velocity with respect to the axial coordinate x ($r=0$) is compared to the theory of Rott [see Eq. (4)]. In (b), (c), and (d), the variation of axial streaming velocity with respect to the transverse coordinate r ($x=L/4$, $L/2$, and $3L/4$) is compared to theory.

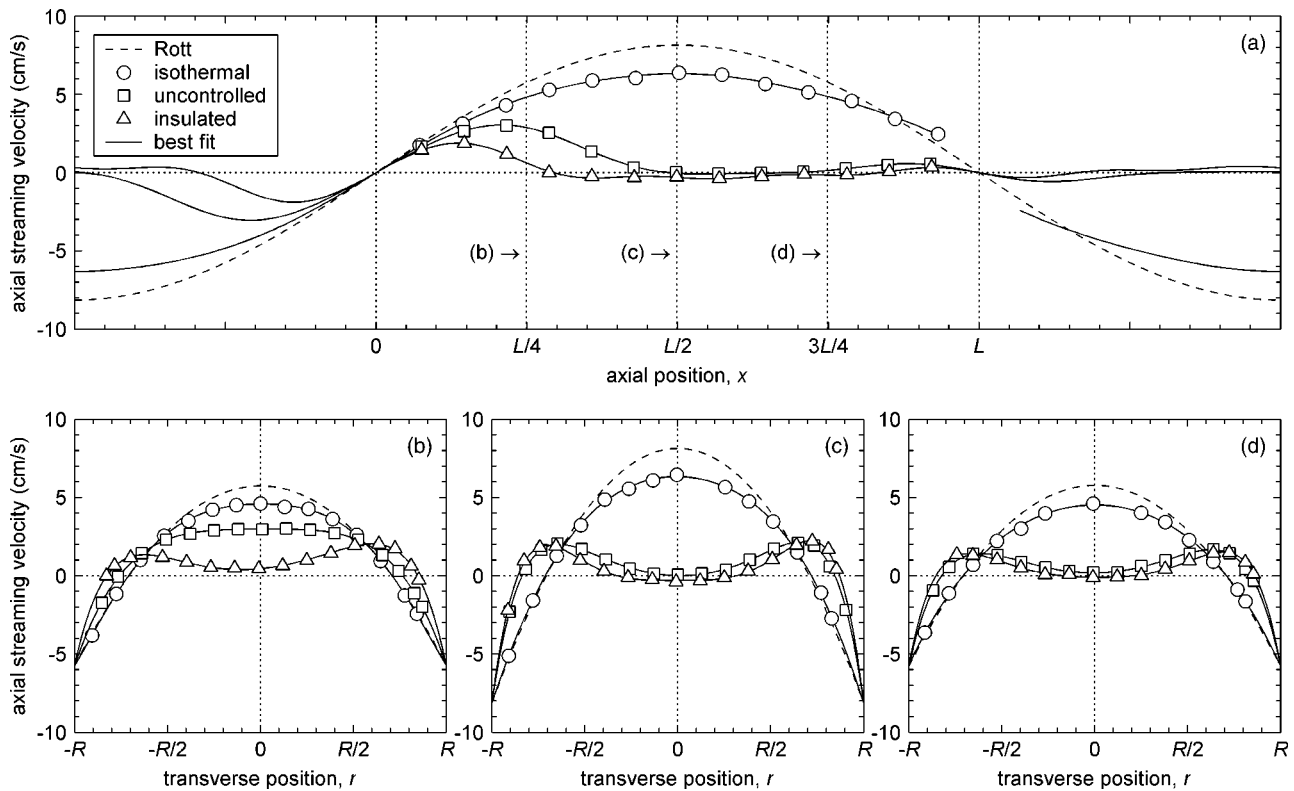


FIG. 8. Axial component of the steady-state streaming-velocity field at a Reynolds number of $Re=40$ ($U_1=8.6$ m/s) for all three boundary conditions. In (a), the variation of axial streaming velocity with respect to the axial coordinate x ($r=0$) is compared to the theory of Rott [see Eq. (4)]. In (b), (c), and (d), the variation of axial streaming velocity with respect to the transverse coordinate r ($x=L/4$, $L/2$, and $3L/4$) is compared to theory.

the resonator walls ($r \approx \pm R$), consistent with the expected cross-sectional average velocities of approximately zero.

Figure 9 shows the axial component of the steady-state streaming-velocity field measured in both the horizontal and vertical planes at a Reynolds number of $Re=40$ with an uncontrolled boundary condition. Measurements in the vertical plane are made using a mirror positioned above the resonator at a 45 deg angle with respect to the horizontal plane. Com-

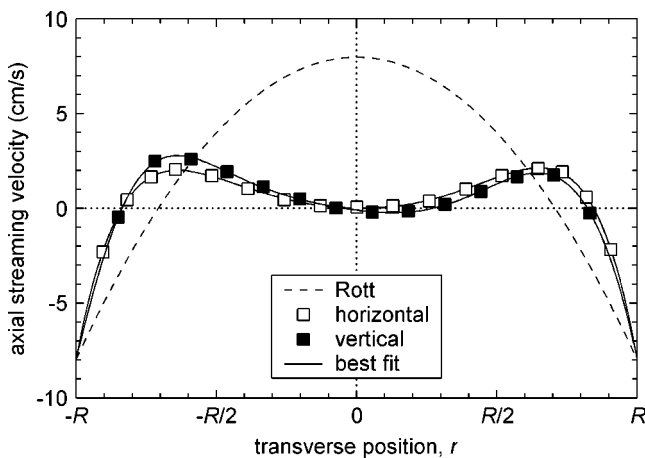


FIG. 9. Axial component of the steady-state streaming-velocity field measured as a function of the transverse coordinate r in both the horizontal and vertical planes ($x=L/2$) at a Reynolds number of $Re=40$ ($U_1=8.5$ m/s) with an uncontrolled boundary condition. For the measurements made in the vertical plane, the position $r=R$ corresponds to the upper wall of the resonator and the position $r=-R$ corresponds to the lower wall. The measurements made in the horizontal plane are taken from Fig. 8(c).

parison of the two curves indicates that the streaming field is slightly asymmetric about the centerline ($r=0$).

C. Decay of streaming

Figure 10 shows the measured decay of streaming when the acoustic field is switched off. For both uncontrolled and insulated boundary conditions, the streaming velocities extrapolated back to time $t=0$ s are in agreement with the steady-state values shown in Fig. 7. The elapsed time of approximately 5 s required for the streaming to completely decay in Fig. 10(a) is in agreement with the time required for the streaming to initially reach its maximum value when the acoustic field is switched on, as measured by Thompson and Atchley.¹ For the larger steady-state axial temperature gradient associated with the insulated boundary condition in Fig. 10(b), the decay curve is more complex.

It is possible to explain the observed behavior in Fig. 10 by supposing the existence of two opposing mechanisms acting on the fluid simultaneously. The first mechanism exerts a force on the fluid in the positive axial direction, has a magnitude that is independent of the axial temperature gradient, and requires approximately 5 s to completely decay. For the smaller temperature gradient associated with the uncontrolled boundary condition in Fig. 10(a), only this first mechanism is observed. The second mechanism exerts a force on the fluid in the negative axial direction, has a magnitude that is a function of the axial temperature gradient, and requires approximately 12 s to completely decay. For the larger temperature gradient associated with the insulated

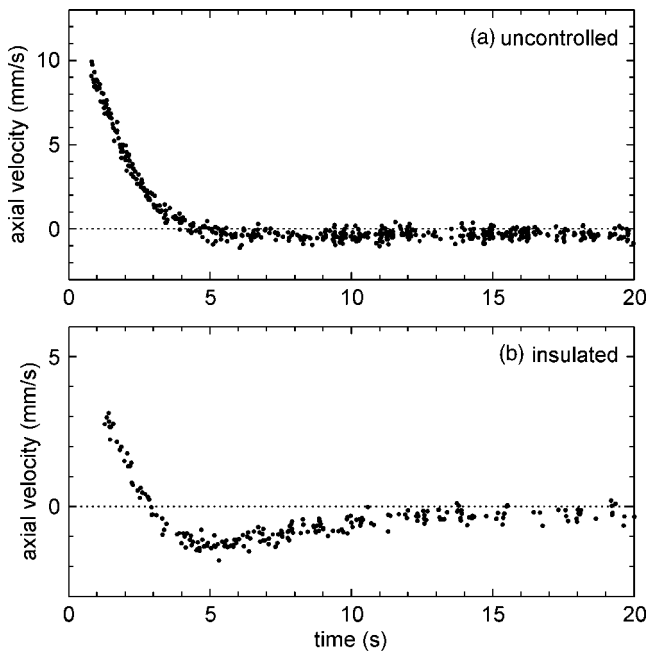


FIG. 10. The axial component of velocity at the center of the streaming cell ($x=L/2$, $r=0$), measured as a function of time after the acoustic field is switched off, for uncontrolled and insulated boundary conditions. The acoustic velocity (not shown) decays rapidly from its steady-state value ($U_1/\sqrt{2}=4.32$ m/s, $Re=20$) to zero within approximately 0.3 s. The subsequent decay of streaming velocity is then measured directly. In (a), the axial temperature difference between the two ends of the streaming cell is approximately 1.0°C (see Fig. 4) and the streaming decays exponentially within approximately 5 s. In (b), the axial temperature difference is approximately 2.3°C and the decay of the streaming is more complex. Note the difference in the scales of the vertical axes.

boundary condition in Fig. 10(b), both mechanisms are observed, causing the streaming velocity to briefly change direction ($t>3$ s) once the first mechanism has decayed sufficiently so that it is no longer dominant. However, this is purely speculative.

D. Ambient convection

After the streaming decays completely in Fig. 10, the ambient convection currents within the resonator are observed to be negligibly small at the center of the streaming cell ($x=L/2$, $r=0$).

For the conditions corresponding to the largest temperature gradient used in the present study ($U_1=8.6$ m/s, $Re=40$, insulated boundary condition), measurements of the streaming decay as a function of the axial and transverse coordinates in the horizontal plane indicate that the ambient convection currents in the horizontal plane are on the order of 1 mm/s—considerably less than the observed steady-state streaming velocities shown in Fig. 8, which are on the order of 80 mm/s near the resonator wall.

The ambient convection currents in the vertical direction (the direction perpendicular to the horizontal plane in which the measurements are made) are observed by rotating the LDA probe by 90° so that its axis of sensitivity points upward instead of sideways. The vertical component of the steady-state fluid velocity is measured as a function of the axial and transverse coordinates in the horizontal plane, and

the time-averaged velocity-signal components are determined. For the above-described conditions ($U_1=8.6$ m/s, $Re=40$, insulated boundary condition), the ambient convection currents in the vertical direction are on the order of 5 mm/s—also less than the observed steady-state axial streaming velocities shown in Fig. 8. The magnitude of the vertical convection is largest in the vicinity of the acoustic-velocity antinode ($x=L$). In this region, the fluid rises near the resonator centerline ($r=0$) and falls near the walls ($r=\pm R$). In the vicinity of the acoustic-velocity node ($x=0$), the magnitude of the vertical convection goes to zero.

V. CONCLUSIONS

The influence of a thermoacoustically induced axial temperature gradient on the axial component of the streaming-velocity field has been observed to be significant (see Sec. IV). In the presence of a relatively small temperature gradient, the measured streaming velocities are not in agreement with the theory of Rott² or with any other available theory (see Secs. II A and II B). The error between the measured and theoretical streaming velocities increases as the temperature gradient is increased. At a Reynolds number of $Re=40$ and with a temperature difference of approximately 1°C or more between the two ends of the streaming cell, the axial streaming along a large section of the centerline ($r=0$) near the acoustic-velocity antinode ($x\approx L$) is completely suppressed.

In the measurements of acoustic streaming reported previously in the literature (see Sec. I), the influence of a temperature gradient on the streaming-velocity field was not addressed. All of the previous measurements made using the PIV technique were limited to the vicinity of the acoustic-velocity node, where the influence of a temperature gradient is least noticeable (see Sec. IV B). Also, if these measurements were carried out shortly after switching on the acoustic field, a significant temperature gradient would not have had enough time to develop. Only Thompson and Atchley⁷ reported a nonparabolic variation of axial streaming velocity with respect to the transverse coordinate r . However, they failed to recognize that this result was due to the presence of a thermoacoustically induced axial temperature gradient, and instead suggested incorrectly that it might be due to the influence of fluid inertia.

Measurements of the axial component of the streaming-velocity field have been made with an isothermal boundary condition imposed in order to observe the influence of fluid inertia on the streaming (see Secs. IV A and IV B). At a Reynolds number of $Re=4$, these measurements are not in agreement with the computational results of Menguy and Gilbert⁸ (see Sec. II B), but rather with the theory of Rott² (see Sec. II A). Measurements made with an isothermal boundary condition at Reynolds numbers of $Re\geq 10$ do show significant discrepancies relative to the theory of Rott, but this effect is less significant than the influence of a temperature gradient on the streaming (see Sec. IV B).

These results suggest that the influence of small temperature gradients needs to be considered carefully when attempting to accurately model the outer streaming-velocity field in a real system. Although Rott,² Olson and Swift,¹⁰ and Bailliet *et al.*¹¹ all assumed a nonzero axial temperature gra-

dient in their analyses (see Sec. II A), their results fail to predict the observed behavior. Perhaps this is because they did not consider the convective heat transport due to the streaming, an effect that Mozurkewich¹⁷ has shown experimentally to be significant.

It may be possible that the heat carried by the streaming is significant enough to affect the spatial distribution of mean temperature within the streaming cell. Perhaps such a change in the mean-temperature distribution could, in turn, affect the streaming-velocity field. In order to investigate this possibility, it would be helpful to measure the mean-temperature distribution within the streaming cell in an unobtrusive manner, so as not to disturb the streaming-velocity field.

As mentioned in Sec. III A, spurious results were observed when measurements of streaming velocity were made shortly after cleaning the inner wall of the resonator using an ammonia-based household glass cleaner. Perhaps the heat transfer between the glass tube and the enclosed air is affected by any residue left on the glass by the cleaner, subsequently affecting the observed streaming motion.

It would be interesting to repeat the experiments described in the present study with the streaming cell oriented vertically instead of horizontally. However, special attention would be needed in order to avoid any thermally induced ambient convection currents that might occur in such a configuration.¹

In addition to investigating experimentally how the streaming-velocity field is affected by the presence of an axial temperature gradient along the resonator wall, it would also be helpful to model the problem mathematically. A simple model might consist of a cylinder of air having two boundary conditions imposed along its surface: (1) a sinusoidally varying axial streaming velocity of the form given by Eq. (4) evaluated for $r = \pm R$, and (2) a sinusoidally varying mean-temperature distribution similar to those observed in the present study (see Fig. 3).

ACKNOWLEDGMENTS

This work has been supported by the Office of Naval Research and The Pennsylvania State University Graduate Program in Acoustics and Applied Research Laboratory Exploratory and Foundational Program. The authors would like

to thank Scott Backhaus, Tetsushi Biwa, Said Boluriaan, Murray Campbell, Andrew Doller, Steven Garrett, Joël Gilbert, Mark Hamilton, Yurii Ilinskii, Robert Keolian, Philip Morris, Matthew Poese, Victor Sparrow, Greg Swift, and Ray Wakeland for numerous helpful discussions during the course of this work.

¹M. W. Thompson and A. A. Atchley, "Simultaneous measurement of acoustic and streaming velocities in a standing wave using laser Doppler anemometry," *J. Acoust. Soc. Am.* **117**, 1828–1838 (2005).

²N. Rott, "The influence of heat conduction on acoustic streaming," *Z. Angew. Math. Phys.* **25**, 417–421 (1974).

³J. P. Sharpe, C. A. Greated, C. Gray, and D. M. Campbell, "The measurement of acoustic streaming using particle image velocimetry," *Acustica* **68**, 168–172 (1989).

⁴M. P. Arroyo and C. A. Greated, "Stereoscopic particle image velocimetry," *Meas. Sci. Technol.* **2**, 1181–1186 (1991).

⁵D. B. Hann and C. A. Greated, "The measurement of flow velocity and acoustic particle velocity using particle-image velocimetry," *Meas. Sci. Technol.* **8**, 1517–1522 (1997).

⁶M. Campbell, J. A. Cosgrove, C. A. Greated, S. Jack, and D. Rockliff, "Review of LDA and PIV applied to the measurement of sound and acoustic streaming," *Opt. Laser Technol.* **32**, 629–639 (2000).

⁷M. W. Thompson and A. A. Atchley, "Measurements of Rayleigh streaming in high-amplitude standing waves," in *Nonlinear Acoustics at the Beginning of the 21st Century*, edited by O. V. Rudenko and O. A. Sapozhnikov (MSU Faculty of Physics, Moscow, 2002), Vol. 1, pp. 183–190.

⁸L. Menguy and J. Gilbert, "Non-linear acoustic streaming accompanying a plane stationary wave in a guide," *Acust. Acta Acust.* **86**, 249–259 (2000).

⁹G. W. Swift, *Thermoacoustics: A Unifying Perspective for Some Engines and Refrigerators* (Acoustical Society of America, Melville, NY, 2002), Chap. 7, pp. 153–204.

¹⁰J. R. Olson and G. W. Swift, "Acoustic streaming in pulse tube refrigerators: Tapered pulse tubes," *Cryogenics* **37**, 769–776 (1997).

¹¹H. Bailliet, V. Gusev, R. Raspet, and R. A. Hiller, "Acoustic streaming in closed thermoacoustic devices," *J. Acoust. Soc. Am.* **110**, 1808–1821 (2001).

¹²Q. Qi, "The effect of compressibility on acoustic streaming near a rigid boundary for a plane traveling wave," *J. Acoust. Soc. Am.* **94**, 1090–1098 (1993).

¹³J. Lighthill, "Acoustic streaming," *J. Sound Vib.* **61**, 391–418 (1978).

¹⁴P. Merkli and H. Thomann, "Thermoacoustic effects in a resonance tube," *J. Fluid Mech.* **70**, 161–177 (1975).

¹⁵G. W. Swift, "Thermoacoustic engines," *J. Acoust. Soc. Am.* **84**, 1145–1180 (1988).

¹⁶R. Sonnenberger, K. Graichen, and P. Erk, "Fourier averaging: A phase-averaging method for periodic flow," *Exp. Fluids* **28**, 217–224 (2000).

¹⁷G. Mozurkewich, "Heat transport by acoustic streaming within a cylindrical resonator," *Appl. Acoust.* **63**, 713–735 (2002).

On the interaction of counterpropagating acoustic waves in resonant rods composed of materials with hysteretic quadratic nonlinearity

Vitalyi Gusev^{a)}

Université du Maine, 72085 Le Mans, Cedex 09, France

(Received 18 May 2004; revised 2 July 2004; accepted 29 December 2004)

An ordinary inhomogeneous integrodifferential equation for the profile of the acoustic wave in a resonant rod, composed of a material with hysteretic quadratic nonlinearity, is derived. It explicitly takes into account the interaction of the counterpropagating acoustic waves. It also incorporates the boundary conditions in the sense that all possible solutions of the equation satisfy the appropriate boundary conditions at the ends of the rod. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1860470]

PACS numbers: 43.25.-x, 43.25.Lj, 43.25.Gf, 43.25.Ed [MFH]

Pages: 1850–1857

I. INTRODUCTION

The interest in theoretical description of vibrations of rods composed of materials with hysteretic quadratic nonlinearity is related to the extended applications of the resonance acoustic phenomenon to the evaluation of polycrystalline metals,^{1–3} ceramics, and rocks.^{4–7} There is an increasing amount of experimental evidence that nonlinear acoustic phenomena are more sensitive to the mesoscopic structure of these materials than the linear ones.^{1,8} Importantly, by using resonators, it is easier to reach acoustic wave amplitudes sufficient for clear manifestation of the nonlinear effects than without acoustic energy accumulation.

The theoretical analysis of the nonlinear acoustic waves in a resonator with hysteretic quadratic nonlinearity might be expected to be more complicated in comparison to a resonator with an important role of “classical” elastic quadratic nonlinearity. An example of the latter, which has been extensively studied, is a closed-ended gas-filled tube.^{9–13} Significant simplification of the theoretical description in this system is due to the absence of such interactions between the counter-propagating waves which lead to the acoustic profile distortion accumulating at distances exceeding the acoustic wavelength. From a physics point of view the forces created by the interaction of the counterpropagating waves in the medium with elastic quadratic nonlinearity (which is even in its physical manifestations) are not moving forces. Because of this they are not synchronous with any of the counter-propagating waves and they are not inducing any accumulating changes in the wave profiles. The possibility to present acoustic field in the resonator with effective elastic quadratic nonlinearity as a linear superposition of two noninteracting nonlinear acoustic waves is an explicit result of Chester’s theory.⁹ It was also clearly stated by Ostrovsky¹⁰ and has been later used by a number of the authors to derive both the equations for the stationary wave profile¹¹ and the nonlinear evolution equations describing the build-up of the nonlinear

oscillations following the start-up of the resonator excitation.^{12,13}

The physics of the acoustic waves interaction in the resonators with hysteretic quadratic nonlinearity is quite different because being quadratic in wave amplitude (similar to elastic quadratic nonlinearity) it is odd in its physical manifestations^{14–17} (in contrast to elastic quadratic nonlinearity, which is even). For example this nonlinearity is predicted to induce self-focusing of the acoustic beams.¹⁴ For the current discussion it is important that the interaction of the counterpropagating waves due to hysteretic quadratic nonlinearity creates the forces, which are synchronous with the acoustic waves. Consequently the acoustic field in the resonant rods cannot be decomposed on two noninteracting nonlinear counterpropagating waves. The theory should explicitly account for the effective interaction between them. Traditionally for the description of the nonlinear standing acoustic waves in hysteretic rods the second-order nonlinear partial differential equations are used.^{1–3} To the best of our knowledge, the description of the nonlinear acoustic wave profile in the resonators with hysteretic quadratic nonlinearity in the form of an ordinary (i.e., not a partial) integrodifferential equation has never been proposed before. Perhaps, this is related to the above-discussed complexity.

Below an ordinary inhomogeneous integrodifferential equation for the profiles of nonlinear acoustic waves in rods with hysteretic quadratic nonlinearity is derived. The analysis of its approximate solutions demonstrates that the interaction of the counterpropagating waves leads to increased role of the hysteretic nonlinearity. The amplitude of oscillations in the resonance peak is lower and the frequency shift of the resonance peak is more pronounced in comparison with the theory where the interaction is neglected. It is demonstrated that these kind of phenomena are also expected in the rods with “classical” elastic cubic nonlinearity, which (similar to hysteretic quadratic nonlinearity) is odd in its physical manifestations, but does not induce nonlinear absorption of sound.

^{a)}Electronic mail: vitali.goussev@univ-lemans.fr

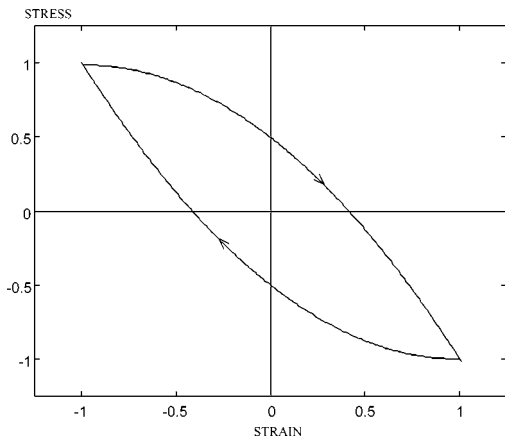


FIG. 1. Presentation in normalized (nondimensional) coordinates of the nonlinear (hysteretic) part of stress–strain relationship for a periodic process with single maximum and single minimum during a period. The arrowheads indicate path direction with increasing time.

II. THEORY

In materials with hysteretic quadratic nonlinearity, the dependence of the dominant strain dependent contribution σ_{hyst} to total stress σ on strain has a form of a cusped loop^{16–18}

$$\sigma_{\text{hyst}} = -hE \left[s^A s + \frac{1}{2} (s^2 - s_A^2) \text{sign} \left(\frac{\partial s}{\partial t} \right) \right]. \quad (1)$$

Here s denotes the strain, s^A denotes strain amplitude, E is the elastic modulus, and $h > 0$ is the nondimensional parameter of the hysteretic quadratic nonlinearity. The stress/strain dependence in Eq. (1) is valid in a periodic strain wave with zero average over a period and a single maximum and a single minimum during a period. In Fig. 1 the hysteretic contribution to stress is presented in nondimensional variables $\sigma'_{\text{hyst}} = \sigma_{\text{hyst}} / (hE)$, $s' = s / s^A$. In accordance with Eq. (1) the stress depends not only on the current value s of the strain but also on the direction of the strain variation (“loading” when $\partial s / \partial t < 0$, “unloading” when $\partial s / \partial t > 0$) and the strain amplitude s^A . The first part of the modulus in Eq. (1) is nonhysteretic (it does not depend on the strain rate $\partial s / \partial t$ at all). This part does not contribute to the nonlinear hysteretic absorption but strongly contributes to the amplitude-dependent sound velocity. The second term in Eq. (1) describes hysteresis in the stress behavior.

It is well established that in the micro-inhomogeneous materials the hysteretic nonlinearity dominates over classical elastic nonlinearity in the stress/strain relationship,^{2,8} and that the kinematic nonlinearity (due to the nonlinear relation between strain and displacement gradient) is also negligible. Due to this, approximating the strain by the displacement gradient $s \approx \partial u / \partial x$, the nonlinear wave equation is presented in the form

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} - c_0^2 \frac{\partial^2 u}{\partial x^2} = & -c_0^2 h \left\{ \left[\left(\frac{\partial u}{\partial x} \right)^A + \frac{\partial u}{\partial x} \text{sign} \left(\frac{\partial^2 u}{\partial x \partial t} \right) \right] \frac{\partial^2 u}{\partial x^2} \right. \\ & \left. + \left[\frac{\partial u}{\partial x} - \left(\frac{\partial u}{\partial x} \right)^A \text{sign} \left(\frac{\partial^2 u}{\partial x \partial t} \right) \right] \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial x} \right)^A \right\}. \end{aligned} \quad (2)$$



FIG. 2. The rod of a finite length L . The motion of the end at $x=L$ is predetermined by the external loading and is assumed to be known in advance. The end at $x=0$ is mechanically free.

In Eq. (2) $c_0 = \sqrt{E / \rho_0}$ denotes the velocity of the infinitely weak acoustic disturbances (the linear sound velocity), ρ_0 is the equilibrium density, and $(\partial u / \partial x)^A$ is the amplitude of the displacement gradient.

The solution of the nonlinear partial differential equation of the second order (2) should satisfy the boundary conditions at the ends of the rod. The geometry of the system under consideration is sketched in Fig. 2. In most of the experiments^{1–7} the external excitation controls the motion of one of the rod ends (the end at $x=L$ in Fig. 2) while the other end (at $x=0$ in Fig. 2) is left mechanically free. The boundary conditions can be written in the form

$$v(x=L, t) = v_0 f(t), \quad (3)$$

$$\rho_0 c_0^2 \partial u / \partial x(x=0, t) + \sigma_{\text{hyst}}(x=0, t) = 0. \quad (4)$$

Here $v = \partial u / \partial t$ denotes the particle velocity, the normalized function $f(t)$ describes the motion of the boundary $x=L$ due to external loading, v_0 is the particle velocity amplitude provided by the external excitation.

It is worth mentioning here that even if the classical elastic quadratic nonlinearity was included in Eq. (2) for generality it would not lead in the system under consideration (Fig. 2) to distortions accumulating at distances exceeding the acoustic wavelength. This is due to mechanically free boundary $x=0$, where the polarity of the incident acoustic wave is inverted in each reflection. Actually if, for example, maximum of v profile moves faster than linear sound in the incident acoustic wave (propagating from $x=L$ to $x=0$) then in the reflected wave (propagating from $x=0$ to $x=L$) it becomes a minimum of v profile moving slower than linear sound. As a result there is no accumulation of the wave profile distortion after each two round-trips of the acoustic wave in the rod, unless the shock front is formed at a single round trip. The latter opportunity will be excluded in the analysis presented in the following. It will be assumed that the wave profile transformations at the scale of the resonator length are weak. Note that in the gas filled resonators with rigid walls^{9–12} the situation is completely different because there is no inversion of acoustic wave polarity in reflection from rigid boundaries. There the elastic quadratic nonlinearity is effective.

The solution u of Eq. (2) can be presented in form of a linear superposition of two interacting counterpropagating waves u_+ and u_- with slowly varying profiles,

$$u = \mu u_+(\mu x, \tau_+ = t - x/c_0) + \mu u_-(\mu x, \tau_- = t + x/c_0). \quad (5)$$

Here the dimensionless symbolic small parameter $\mu \ll 1$ characterizes weakness of the acoustic disturbances and slowness of their nonlinear transformation in propagation.

The accumulation of the wave profile distortion is assumed to take place at distances significantly exceeding the acoustic wavelength λ (which is of the order of the resonator length L). When Eq. (5) is substituted in Eq. (2) the terms of the first order in μ are mutually cancelled. If only the terms of the second order in μ are retained, then Eq. (2) takes the form

$$\begin{aligned} \frac{\partial}{\partial x}(v_+ - v_-) + \frac{h}{2c_0^2} & \left\{ \left[(v_+ - v_-)^A + (v_+ - v_-) \right. \right. \\ & \times \text{sign} \left(\frac{\partial v_+}{\partial \tau_+} - \frac{\partial v_-}{\partial \tau_-} \right) \left. \left. \left(\frac{\partial v_+}{\partial \tau_+} + \frac{\partial v_-}{\partial \tau_-} \right) \right] \right. \\ & + 2 \left[(v_+ - v_-) - (v_+ - v_-)^A \text{sign} \left(\frac{\partial v_+}{\partial \tau_+} - \frac{\partial v_-}{\partial \tau_-} \right) \right] \\ & \left. \times \frac{\partial}{\partial (\tau_+ - \tau_-)} (v_+ - v_-)^A \right\} = 0, \end{aligned} \quad (6)$$

where $v_{\pm} = \partial u_{\pm} / \partial t = \partial u_{\pm} / \partial \tau_{\pm}$ are particle velocities in the counterpropagating waves. Importantly, although the amplitudes of both waves are slowly varying in space [see Eq. (5)], the amplitude of the total strain $(\partial u / \partial x)^A$ (which is a function of x only) does not necessarily vary slowly in space. In other words the derivative $\partial / \partial x$ applied in Eq. (2) to the distribution of the total strain amplitude $(\partial u / \partial x)^A$ does not necessarily diminishes the order of this term by μ factor. The term $\partial(\partial u / \partial x)^A / \partial x$ might be of the first order in μ and not of the second order as it is in the case of the uni-directional (co-propagating) waves.^{14,15,17,18} A simple illustrative example is the following. If we have only a right-propagating wave of a constant amplitude A ($u = u_+(\tau_+) = A \sin[\omega(t - x/c_0)]$), then clearly $(\partial u / \partial x)^A = A \omega / c_0 = \text{const}$ and $\partial(\partial u / \partial x)^A / \partial x = 0$. But if we have two counter-propagating waves of equal and constant amplitudes ($u = u_+(\tau_+) + u_-(\tau_-) = A \sin[\omega(t - x/c_0)] + A \sin[\omega(t + x/c_0)] = 2A \sin(\omega t) \cos(\omega x / c_0)$), then $(\partial u / \partial x)^A = 2A(\omega / c_0) \sin(\omega x / c_0) \neq \text{const}$ and $\partial(\partial u / \partial x)^A / \partial x = 2A(\omega / c_0)^2 \text{sign}[\sin(\omega x / c_0)] \cos(\omega x / c_0) \neq 0$. The latter is an obvious result of the wave interference, providing spatial modulation of the total strain

amplitude at the scale of the acoustic wavelength even without any variation in space of the amplitudes of each of the counter-propagating waves (i.e., even when $\mu = 0$). Due to the fact that $\partial(\partial u / \partial x)^A / \partial x$ is in general of the order of μ terms in Eq. (2), which include this spatial derivative, contribute to the evolution equation (6) where all the terms are of the second order in μ . To distinguish this derivative from the derivative over the slow coordinate μx the former is represented in Eq. (6) by the form $\partial / \partial x = -(2/c_0) \partial / \partial (\tau_+ - \tau_-)$, which just takes into account that $\tau_+ - \tau_- = -2x/c_0$.

Equation (6) formally contains two unknowns (v_+ and v_-) and three variables (fast variables τ_+ , τ_- and slow variable μx). However the variables are not independent: there is an evident relation ($\tau_- = \tau_+ + 2x/c_0$) between them. Due to this relation in the coordinate system $(\mu x, \tau_+)$, accompanying the wave propagating to the right, the function $v_-(\mu x, \tau_-) = v_-(\mu x, \tau_+ + 2x/c_0)$ is a fast varying function of x co-ordinate. This is due to the explicit dependence of the second (fast) argument on x in this system of coordinates. Consequently the term $\partial v_- / \partial x$ in Eq. (6) can be eliminated by averaging in this accompanying coordinate system over the wavelength λ in space (or equivalently over acoustic period T in time). When applying to Eq. (6) the operator $(1/\lambda) \int_{x_0 - \lambda/2}^{x_0 + \lambda/2} (\dots) dx$, it should be taken into account that the first (slow) argument μx of the functions does not provide any significant distortion at the spatial scale λ , and that odd nonlinearity cannot induce average fields. Because of this

$$\begin{aligned} \frac{1}{\lambda} \int_{x_0 - \lambda/2}^{x_0 + \lambda/2} v_- \left(\mu x', \tau_+ + \frac{2x'}{c_0} \right) dx' \\ = \frac{1}{T} \int_{-T/2}^{T/2} v_-(\mu x, \tau_+ + 2t') dt' \\ = \frac{1}{2T} \int_{-T}^T v_-(\mu x, \tau_+ + t') dt' = 0. \end{aligned}$$

Accordingly, Eq. (6) takes the form

$$\begin{aligned} \frac{\partial v_+}{\partial x} = -\frac{h}{2c_0^2} \frac{1}{2T} \int_{-T}^T & \left\{ \left[v_+(\mu x, \tau_+) - v_-(\mu x, \tau_+ + t') \right]^A + \left[v_+(\mu x, \tau_+) - v_-(\mu x, \tau_+ + t') \right] \right. \\ & \times \text{sign} \left[\frac{\partial v_+(\mu x, \tau_+)}{\partial \tau_+} - \frac{\partial v_-(\mu x, \tau_+ + t')}{\partial \tau_-} \right] \left[\frac{\partial v_+(\mu x, \tau_+)}{\partial \tau_+} + \frac{\partial v_-(\mu x, \tau_+ + t')}{\partial \tau_-} \right] \\ & - 2 \left\{ \left[v_+(\mu x, \tau_+) - v_-(\mu x, \tau_+ + t') \right] - \left[v_+(\mu x, \tau_+) - v_-(\mu x, \tau_+ + t') \right]^A \right. \\ & \left. \times \text{sign} \left[\frac{\partial v_+(\mu x, \tau_+)}{\partial \tau_+} - \frac{\partial v_-(\mu x, \tau_+ + t')}{\partial \tau_-} \right] \right\} \frac{\partial}{\partial t'} \left[v_+(\mu x, \tau_+) - v_-(\mu x, \tau_+ + t') \right]^A \Bigg\} dt' \\ = -\frac{h}{2c_0^2} \frac{1}{2T} \int_{-T}^T & \Phi[v_+(\mu x, \tau_+), v_-(\mu x, \tau_+ + t')] dt'. \end{aligned} \quad (7)$$

Obviously the application of the same averaging procedure in the coordinate system $(\mu x, \tau_-)$, accompanying the wave propagating to the left, results in

$$\begin{aligned}
\frac{\partial v_-}{\partial x} &= \frac{h}{2c_0^2} \frac{1}{2T} \int_{-T}^T \left\{ \left[v_+(\mu x, \tau_- - t') - v_-(\mu x, \tau_-) \right]^A + \left[v_+(\mu x, \tau_- - t') - v_-(\mu x, \tau_-) \right] \right. \\
&\quad \times \text{sign} \left[\frac{\partial v_+(\mu x, \tau_- - t')}{\partial \tau_+} - \frac{\partial v_-(\mu x, \tau_-)}{\partial \tau_-} \right] \left. \left[\frac{\partial v_+(\mu x, \tau_- - t')}{\partial \tau_+} + \frac{\partial v_-(\mu x, \tau_-)}{\partial \tau_-} \right] \right. \\
&\quad - 2 \left\{ \left[v_+(\mu x, \tau_- - t') - v_-(\mu x, \tau_-) \right] - \left[v_+(\mu x, \tau_- - t') - v_-(\mu x, \tau_-) \right]^A \right. \\
&\quad \times \text{sign} \left[\frac{\partial v_+(\mu x, \tau_- - t')}{\partial \tau_+} - \frac{\partial v_-(\mu x, \tau_-)}{\partial \tau_-} \right] \left. \left. \frac{\partial}{\partial t'} \left[v_+(\mu x, \tau_- - t') - v_-(\mu x, \tau_-) \right]^A \right\} dt' \right. \\
&\equiv \frac{h}{2c_0^2} \frac{1}{2T} \int_{-T}^T \Phi[v_+(\mu x, \tau_- - t'), v_-(\mu x, \tau_-)] dt'. \tag{8}
\end{aligned}$$

In Eqs. (7) and (8) compact notations for the right-hand sides (r.h.s.) has been introduced. These equations provide the system of coupled partial integrodifferential equations for v_+ and v_- . Each of these equations is only in two independent variables. The averaging procedure applied to nonlinear terms retains in Eq. (7) and in Eq. (8) only those components of the nonlinear force which are synchronous with the right-propagating wave and left-propagating wave, respectively.

In the resonator (Fig. 2) the solution of Eqs. (7) and (8) should be subjected to boundary conditions (3) and (4). It is useful to incorporate the boundary conditions in the equations. This can be done by applying the method of successful approximations developed by Chester.^{9,12} Chester noticed that when Eqs. (7) and (8) are integrated over x at spatial scales of the order of the resonator wavelength L it is possible to neglect the dependence of v_+ and v_- on slow coordinate μx in the r.h.s. of these equations. This is due to assumed weak distortion of the wave profiles at the distances of the order of the acoustic wavelength. Consequently, the unknowns $v_{\pm}(\mu x, \tau_{\pm})$ on the r.h.s. of Eqs. (7) and (8) can be replaced by the profiles $v_{\pm}(\tau_{\pm})$ explicitly independent of space coordinate,

$$\begin{aligned}
\frac{\partial v_+}{\partial x} &= -\frac{h}{2c_0^2} \frac{1}{2T} \int_{-T}^T \Phi[v_+(\tau_+), v_-(\tau_+ + t')] dt', \\
\frac{\partial v_-}{\partial x} &= \frac{h}{2c_0^2} \frac{1}{2T} \int_{-T}^T \Phi[v_+(\tau_- - t'), v_-(\tau_-)] dt'.
\end{aligned}$$

The integration of these equations results in

$$\begin{aligned}
v_+(\mu x, \tau_+) &= g_+(\tau_+) - \frac{h}{2c_0^2} \left\{ \frac{1}{2T} \right. \\
&\quad \times \left. \int_{-T}^T \Phi[v_+(\tau_+), v_-(\tau_+ + t')] dt' \right\} x, \tag{9}
\end{aligned}$$

$$\begin{aligned}
v_-(\mu x, \tau_-) &= g_-(\tau_-) + \frac{h}{2c_0^2} \left\{ \frac{1}{2T} \right. \\
&\quad \times \left. \int_{-T}^T \Phi[v_+(\tau_- - t'), v_-(\tau_-)] dt' \right\} x, \tag{10}
\end{aligned}$$

where $g_{\pm}(\tau_{\pm})$ are unknown functions (of a single argument each) to be determined from the boundary conditions. Actu-

ally from Eqs. (9) and (10) it follows that $g_{\pm}(t)$ are the particle velocities in the counterpropagating waves at the mechanically free end $x=0$,

$$\begin{aligned}
v_+(\mu x, \tau_+)^{x=0} &= v_+(0, t) = g_+(t), \\
v_-(\mu x, \tau_-)^{x=0} &= v_-(0, t) = g_-(t).
\end{aligned}$$

They are related by the boundary condition (4). Importantly only the relation of the first order in $\mu \ll 1$ should be retained in Eq. (4),

$$v_+(0, t) - v_-(0, t) = 0. \tag{11}$$

It is verified that all the higher order corrections to Eq. (11) [like those coming from the nonlinear stress σ_{hyst} in Eq. (4)] provide contributions to the equation for the acoustic wave profile of the orders higher than μ^2 . In accordance with Eq. (11) $g_-(t) = g_+(t) \equiv g(t)$ and the solutions (9) and (10) take the form

$$\begin{aligned}
v_+(\mu x, \tau_+) &= g(\tau_+) - \frac{h}{2c_0^2} \left\{ \frac{1}{2T} \int_{-T}^T \Phi[g(\tau_+), g(\tau_+ + t')] dt' \right\} x, \tag{12}
\end{aligned}$$

$$\begin{aligned}
v_-(\mu x, \tau_-) &= g(\tau_-) + \frac{h}{2c_0^2} \left\{ \frac{1}{2T} \int_{-T}^T \Phi[g(\tau_- - t'), g(\tau_-)] dt' \right\} x. \tag{13}
\end{aligned}$$

Here the space-independent profiles v_{\pm} on the r.h.s. are approximated by their form at the mechanically free boundary. The substitution of the solutions (12) and (13) at the other end of the rod (that is evaluated at $x=L$) into the boundary condition (3) leads to the closed form integrodifferential equation for the wave profile $g(t)$,

$$g(t-L/c_0) + g(t+L/c_0) - \frac{h}{2c_0^2} \left\{ \frac{1}{2T} \int_{-T}^T \Phi[g(t-L/c_0), g(t-L/c_0+t')] - \Phi[g(t+L/c_0-t'), g(t+L/c_0)] \right\} dt' = v_0 f(t). \quad (14)$$

In Eq. (14) both the nonlinear term (in the form of the integral) and the external loading (r.h.s.) are of the second order in $\mu \ll 1$. Consequently, in the first order the relation $g(t-L/c_0) = -g(t+L/c_0)$ is valid. The latter relation is used to simplify the nonlinear term in Eq. (14). Using also the symmetry and periodicity properties of the operator Φ , Eq. (14) is reduced to the form

$$g(t-L/c_0) + g(t+L/c_0) - \frac{hL}{c_0^2} \left\{ \frac{1}{T} \int_{-T/2}^{T/2} \Phi[g(t-L/c_0), g(t-L/c_0+t')] \right\} dt' = v_0 f(t).$$

Introducing the notations $t-L/c_0 \equiv \tau$ and $f(\tau+L/c_0) \equiv \bar{f}(\tau)$ it can be rewritten as

$$g(\tau) + g(\tau+2L/c_0) - \frac{hL}{c_0^2} \left\{ \frac{1}{T} \int_{-T/2}^{T/2} \Phi[g(\tau), g(\tau+t')] \right\} dt' = v_0 \bar{f}(\tau). \quad (15)$$

Considering periodic harmonic excitation, the function $\bar{f}(\tau)$ is chosen in the form $\bar{f}(\tau) = \cos(\omega\tau)$. Then the smallness

($\propto \mu^2$) of the sum $g(\tau) + g(\tau+2L/c_0)$ in Eq. (15) requires a small deviation of $2L/c_0$ from the half-period $T/2 = \pi/\omega$ of the excitation ($2L/c_0 = T/2 + 2\Delta L/c_0$, $\Delta L/L \propto \mu \ll 1$) and the symmetry $g(\tau+T/2) \equiv -g(\tau)$ of the wave profile. Because of small deviation ΔL of rod length from the linear resonance length $L = Tc_0/4$ (for the fundamental $\lambda/4$ resonance con-

sidered in the following) the function $g(\tau+2L/c_0)$ can be approximated by Taylor expansion as $g(\tau+2L/c_0) \equiv -g(\tau) - \partial g / \partial \tau (\tau) (2\Delta L/c_0)$. With the help of this relation, Eq. (15) is rewritten

$$\frac{2\Delta L}{c_0} \frac{\partial g}{\partial \tau} + \frac{hL}{c_0^2} \frac{1}{T} \int_{-T/2}^{T/2} \Phi[g(\tau), g(\tau+t')] dt' = v_0 \cos(\omega\tau), \quad (16)$$

where all the terms are of the same (μ^2) order of smallness.

For the subsequent analysis it is useful to introduce dimensionless time $\theta = \omega\tau$ and dimensionless particle velocity $V = g/v_{NL}$, where $v_{NL} = c_0 \sqrt{v_0/h\omega L} \equiv \sqrt{2c_0 v_0/\pi h}$ is the characteristic amplitude of acoustic wave in the hysteretic rod. Note that v_{NL} is proportional to square root of the excitation velocity ($v_{NL} \propto \sqrt{v_0}$) as is expected for the quadratic nonlinearity.^{10,19,20} The dimensionless form of Eq. (15) is

$$\begin{aligned} \delta \frac{\partial V}{\partial \theta} + \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ [V(\theta) - V(\theta+\theta')]^A \right. \\ \left. + [V(\theta) - V(\theta+\theta')] \text{sign} \left[\frac{\partial V(\theta)}{\partial \theta} - \frac{\partial V(\theta+\theta')}{\partial \theta} \right] \right\} \\ \times \left[\frac{\partial V(\theta)}{\partial \theta} + \frac{\partial V(\theta+\theta')}{\partial \theta} \right] - 2 \left\{ [V(\theta) - V(\theta+\theta')] \right. \\ \left. - [V(\theta) - V(\theta+\theta')]^A \text{sign} \left[\frac{\partial V(\theta)}{\partial \theta} - \frac{\partial V(\theta+\theta')}{\partial \theta} \right] \right\} \\ \times \frac{\partial}{\partial \theta'} [V(\theta) - V(\theta+\theta')]^A \Big|_{d\theta'} = \cos \theta. \end{aligned} \quad (17)$$

Equation (17) belongs to the class of the equations describing excitation of the nonlinear waves by moving quasisynchronous sources.^{19,20} In accordance with Eq. (17) the profiles of the waves traveling in the rod depend on a single dimensionless parameter δ , which characterizes the detuning of the system from the linear resonance. It can be expressed in terms of a deviation $\Delta\omega = \omega - \omega_0$ of the excitation frequency ω from linear resonance frequency $\omega_0 = \pi c_0/2L$ as $\delta = \Delta\omega/\omega_{NL}$, where $\omega_{NL} = (\omega_0/\pi)(v_0/v_{NL}) = (\sqrt{hv_0/2\pi c_0})\omega_0$ is the characteristic shift of the resonance frequency. The shift of the resonance by hysteretic quadratic nonlinearity is expected due to odd character of its physical manifestations. For comparison, the elastic quadratic nonlinearity (which has even character) does not cause any shift of the resonance in closed tubes.^{9,11,12} The above-developed theory clearly predicts that the shift of the resonance in vibrating rod is proportional to the wave amplitude because from $v_{NL} \propto \sqrt{v_0}$ and $\omega_{NL} \propto \sqrt{v_0}$ it follows that $\omega_{NL} \propto v_{NL}$. This theoretical result correlates with numerous experimental observations.^{1,2,4,6}

The derived inhomogeneous ordinary integrodifferential equation (17) is the primary result of the developed theory. It takes into account the influence on the profile of the acoustic wave of its interaction with the counterpropagating wave. These are the profiles $V(\theta+\theta')$ and $\partial V(\theta+\theta')/\partial \theta$ that might be identified in Eq. (17) with the counterpropagating wave. From a physics point of view the closed form single integrodifferential equation is obtainable due to the existing relation $[g_+(t) = g_-(t)]$ between the profiles of the counterpropagating waves in the resonator. Qualitatively speaking, in the resonator a traveling wave interacts with itself, but delayed in time and propagating in the opposite direction.

Unfortunately currently the precise solutions of Eq. (17) are not available. To find an approximate solution at fundamental excitation frequency ω the excitation of the higher harmonics can be neglected in first approximation. A similar approach has been applied earlier^{1,2} to solve directly the nonlinear partial differential equations of the type of Eq. (2). First, the solution of the assumed form $V = V_A \sin(\theta+\varphi)$ is substituted in Eq. (17) and then cosine and sine Fourier amplitudes of the different terms of the equation are balanced separately. This results in the system of equations for the amplitude V_A and the phase φ of the acoustic wave,

$$\delta V_A + 2aV_A^2 = \cos \varphi, \quad a^2 V_A^2 = \sin \varphi, \quad (18)$$

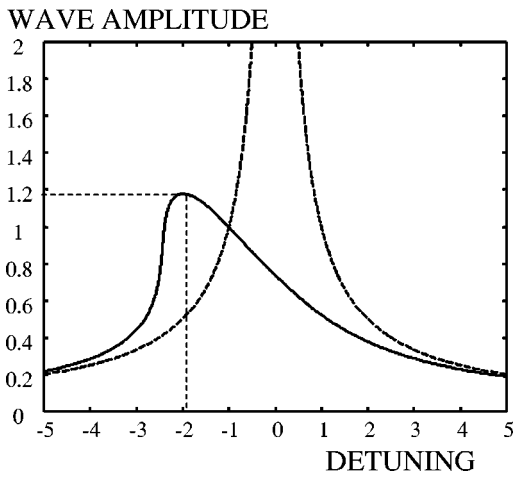


FIG. 3. Theoretical dependence of the dimensionless amplitude V_A of the traveling acoustic wave in the rod on the dimensionless detuning δ of the excitation frequency from the linear resonance condition (corresponding to $\delta=0$). The dashed curves represent linear resonance. The continuous curve depicts the saturation of the wave amplitude and the shift of the resonance peak due to hysteretic quadratic nonlinearity of material composing the rod.

where the constant $a=8/(3\pi)$. Excluding phase from Eq. (18) the description of the resonance curve $V_A=V_A(\delta)$ is obtained in the implicit form

$$\delta = \frac{1}{V_A} [-2aV_A^2 \pm \sqrt{1-a^4V_A^4}]. \quad (19)$$

In accordance with Eq. (19) the maximum amplitude $V_A^{\max}=1/a \approx 1.18$ is achieved for the detuning $\delta = \delta_{\max} = -2$. It is suitable to rewrite Eq. (19) as

$$\delta = \frac{1}{V_A} [-2aV_A^2 + \sqrt{1-a^4V_A^4} \text{sign}(\delta+2)]. \quad (20)$$

The resonance curve (20) is presented in Fig. 3, where for comparison the linear resonance curve $V_A=1/|\delta|$ [when the nonlinear term in Eq. (17) is neglected] is also depicted. Note that diminishing of the resonance frequency ($\delta_{\max} < 0$) predicted by the solution (20) (see Fig. 3) is in accordance with experimental observations, confirming softening of materials with hysteretic quadratic nonlinearity with increased amplitude of the acoustic loading.^{1-4,6-8}

III. DISCUSSION

The derived solution for the resonance peak ($V_A^{\max} \approx 1.18, \delta_{\max} \approx -2$) takes into account the interaction of the counterpropagating waves. If the counterpropagating wave is omitted in Eq. (17) this equation simplifies

$$\delta(\partial V/\partial \theta) + [V^A - V \text{sign}(\partial V/\partial \theta)](\partial V/\partial \theta) = \cos \theta. \quad (21)$$

The solution of this equation by the same mathematical method leads to

$$\delta = \frac{1}{V_A} \left[-V_A^2 + \sqrt{1 - \frac{1}{4}a^2V_A^4 \text{sign}(\delta + \sqrt{2/a})} \right].$$

Consequently, when the interaction of the counterpropagating waves in the rod is neglected the maximum amplitude $\bar{V}_A^{\max} = \sqrt{2/a} \approx 1.53$ is higher and the resonance peak shift

$\bar{\delta}_{\max} = -\sqrt{2/a} \approx -1.53$ is smaller than in the case when the interaction is taken into account. This means that the interaction of the counterpropagating waves amplifies the manifestation of the hysteretic quadratic nonlinearity. If one imagines this interaction to be switched off, then the vibration amplitude will increase and the resonance shift will diminish. Thus the interaction of the counterpropagating waves additionally absorbs each of the waves and makes the material softer.

It is important to evaluate the role in the analyzed phenomenon of the amplitude grating [the term proportional to $\partial(\partial u/\partial x)^A/\partial x$ in Eq. (2)]. If this term is neglected then the nonlinearity of the medium in the wave equation will be accounted for only by the modified local elastic modulus. Actually the coefficient on the right-hand side of Eq. (2) in front of the term $\partial^2 u/\partial x^2$ describes the so-called “bow-tie” contribution to modulus in materials with hysteretic quadratic nonlinearity.^{8,21,22} If only this nonlinear term is retained in Eq. (2) then the solution for the resonance curve will be modified for

$$\delta = \frac{1}{V_A} \left[-aV_A^2 + \sqrt{1 - \frac{1}{4}a^4V_A^4 \text{sign}(\delta + \sqrt{2})} \right],$$

with the maximum amplitude $(V_A^{\max})' = \sqrt{2}/a \approx 1.66$ for the detuning $(\delta_{\max})' = -\sqrt{2} \approx -1.41$. Note that the amplitude is larger (and the shift of the resonance is smaller) not just in comparison with the case where amplitude grating (interference pattern) is taken into account, but also in comparison with the case where the interaction of the counterpropagating waves is switched off. Consequently local variation of the material modulus leads separately (i.e., when separated from the amplitude grating) to the tendency in nonlinear phenomena, which is just opposite to the one predicted above. If only local modulus modifications are taken into account then the interaction of the counterpropagating waves partially suppresses the manifestation of the hysteretic quadratic nonlinearity. If one imagines that only the nonlinear interaction of counterpropagating waves through the local hysteresis modulus is switched on, then the wave amplitude will increase and the resonance shift will diminish. This indicates, in particular, that the interaction of the counterpropagating waves through the bow-tie contribution to modulus may amplify the waves. The effect of the amplification of a weak acoustic wave by a strong counterpropagating acoustic wave in materials with “bow-tie” hysteretic contribution to modulus was predicted earlier in Refs. 21 and 22. The above-presented theory clearly demonstrates that in the case of the counterpropagating waves of comparable amplitudes there is important contribution to nonlinear processes originating from the amplitude grating and that the contribution from the amplitude grating overcompensates for the amplification effects related to local hysteretic modulus. This indicates that the amplitude grating might play a role in the case of the counterpropagating waves of significantly different amplitudes as well, and that the theory developed in Refs. 21 and 22 should be revised.

It is interesting to compare the above-mentioned predictions with those for the rods with “classical” elastic cubic

nonlinearity. A properly normalized equation neglecting interaction of the counterpropagating waves is expected to have the form

$$\delta(\partial V/\partial\theta) + V^2(\partial V/\partial\theta) = \cos\theta. \quad (22)$$

Surely v_{NL} and ω_{NL} for the normalization of Eq. (22) are different from those in Eq. (21). The solution of Eq. (22) is

$$\delta = \frac{1}{V_A} \left[\frac{V_A^3}{3} \pm 1 \right]. \quad (23)$$

If the interaction of the counterpropagating waves is taken into account then Eq. (22) is expected to take the form

$$\delta \frac{\partial V}{\partial\theta} + \frac{1}{2\pi} \int_{-\pi}^{\pi} [V(\theta) - V(\theta + \theta')]^2 \times \left[\frac{\partial V(\theta)}{\partial\theta} + \frac{\partial V(\theta + \theta')}{\partial\theta} \right] d\theta' = \cos\theta. \quad (24)$$

Please compare the transition from Eq. (24) to Eq. (22) with the one from Eq. (17) to Eq. (21). The solution of Eq. (24) is

$$\delta = \frac{1}{V_A} \left[\frac{3V_A^3}{4} \pm 1 \right]. \quad (25)$$

Comparison of Eqs. (25) and (23) demonstrates that the interaction of the counterpropagating waves increases the role of the elastic cubic nonlinearity (the coefficient in front of V_A^3 increases) similar to the case of the hysteretic quadratic nonlinearity. The important difference between the two cases is, however, in the absence of the nonlinear absorption in materials with elastic cubic nonlinearity and in different dependence of the frequency shift on the wave amplitude.

It is instructive to compare the approximate solution $\bar{V}_A^{\max} = \sqrt{3\pi/4} \approx 1.53$, $\bar{\delta}_{\max} = -\sqrt{3\pi/4} \approx -1.53$ of Eq. (21) with its exact solution $\bar{V}_A^{\max} = \sqrt{2} \approx 1.41$, $\bar{\delta}_{\max} = -\sqrt{2} \approx -1.41$, which is available.²³ This comparison demonstrates that neglecting the inverse influence of the generated higher harmonics $[\omega_n = (2n+1)\omega, n=1,2,\dots]$ on the fundamental frequency ω leads only to a slight ($\approx 10\%$) overestimation of both maximum vibration amplitude and the resonance frequency shift. Similar overestimation might be expected in the obtained approximate solution of Eq. (17) in comparison with its exact solution.

Finally it is worth mentioning that the derived Eq. (17) can be extended, if necessary, to account for the effects of weak linear sound velocity dispersion^{24,25} and/or weak linear absorption^{2,9,11,26} by adding corresponding terms in the equation. In the case of weak linear dispersion the corresponding term is proportional to $\partial^3 V/\partial\theta^3$, for example.^{24,25}

IV. CONCLUSIONS

An ordinary inhomogeneous integrodifferential equation for the profile of the acoustic wave in a resonant rod, composed of a material with hysteretic quadratic nonlinearity, is derived. This equation incorporates the boundary conditions in the sense that all possible solutions of this equation satisfy the necessary boundary conditions "automatically." It explicitly takes into account the interaction of the counterpropagating acoustic waves. It is demonstrated that this in-

teraction amplifies the nonlinear phenomena. In particular the maximum vibration amplitude predicted by this theory is lower and the shift of the resonance peak is larger in comparison with the case when the interaction of the counterpropagating waves is neglected. The interaction due to hysteretic quadratic nonlinearity suppresses both counterpropagating waves and also enhances softening of the material with increasing wave amplitude.

ACKNOWLEDGMENT

The author is grateful for fruitful and stimulating discussions of the obtained results to Professor L. A. Ostrovskii, who appeared to be working in parallel on this problem.

- ¹ V. E. Nazarov, L. A. Ostrovskii, I. A. Soustova, and A. M. Sutin, "Anomalous acoustic nonlinearity in metals," *Akust. Zh.* **34**, 491–499 (1988) [English transl.: *Sov. Phys. Acoust.* **34**, 284–289 (1988)].
- ² V. E. Nazarov, L. A. Ostrovskii, I. A. Soustova, and A. M. Sutin, "Nonlinear acoustics of micro-inhomogeneous media," *Phys. Earth Planet. Inter.* **50**, 65–73 (1988).
- ³ V. E. Nazarov and A. B. Kolpakov, "Experimental investigation of nonlinear acoustic phenomena in polycrystalline zinc," *J. Acoust. Soc. Am.* **107**, 1915–1921 (2000).
- ⁴ P. A. Johnson, B. Zinszner, and P. N. J. Rasolofosaon, "Resonance and elastic nonlinear phenomena in rock," *J. Geophys. Res.*, **101**, 11553–11564 (1996).
- ⁵ R. A. Guyer, K. R. McCall, and K. Van Den Abeele, "Slow elastic dynamics in a resonant bar of rock," *Geophys. Res. Lett.* **25**, 1585–1588 (1998).
- ⁶ R. A. Guyer, J. TenCate, and P. Johnson, "Hysteresis and the dynamic elasticity of consolidated granular materials," *Phys. Rev. Lett.* **82**, 3280–3283 (1999).
- ⁷ V. E. Nazarov, A. V. Radostin, and I. A. Soustova, "Effect of an intense sound wave on the acoustic properties of a sandstone bar resonator. Experiment," *Akust. Zh.* **48**, 85–90 (2000) [English transl.: *Acoust. Phys.* **48**, 76–80 (2000)].
- ⁸ R. A. Guyer and P. A. Johnson, "Nonlinear mesoscopic elasticity: Evidence for a new class of materials," *Phys. Today* **52**, 30–35 (1999).
- ⁹ W. Chester, "Resonant oscillations in closed tubes," *J. Fluid Mech.* **18**, 44–64 (1964).
- ¹⁰ L. A. Ostrovskii, "Discontinuous oscillations in acoustic resonator," *Akust. Zh.* **20**, 140–142 (1974) [English transl.: *Sov. Phys. Acoust.* **20**, 88–89 (1974)].
- ¹¹ V. V. Kaner, O. V. Rudenko, and R. V. Khokhlov, "Theory of nonlinear oscillations in acoustic resonators," *Akust. Zh.* **23**, 756–765 (1977) [English transl.: *Sov. Phys. Acoust.* **23**, 432–437 (1977)].
- ¹² V. E. Gusev, "Build-up of forced oscillations in acoustic resonator," *Akust. Zh.* **30**, 204–212 (1984) [*Sov. Phys. Acoust.* **30**, 121–125 (1984)].
- ¹³ V. E. Gusev, H. Bailliet, P. Lotton, and M. Bruneau, "Asymptotic theory of nonlinear acoustic waves in a thermoacoustic prime-mover," *Acust. Acta Acust.* **86**, 25–38 (2000).
- ¹⁴ V. Gusev, C. Glorieux, W. Lauriks, and J. Thoen, "Nonlinear bulk and surface shear acoustic waves in materials with hysteresis and end-point memory," *Phys. Lett. A* **232**, 77–86 (1997).
- ¹⁵ V. Gusev, "Self-modulation instability in materials with hysteretic quadratic nonlinearity," *Wave Motion* **33**, 145–153 (2001).
- ¹⁶ R. A. Guyer, K. P. McCall, and G. N. Boitnott, "Hysteresis, discrete memory, and nonlinear wave propagation in rock: A new paradigm," *Phys. Rev. Lett.* **74**, 3491–3494 (1995).
- ¹⁷ V. E. Gusev, W. Lauriks, and J. Thoen, "Dispersion of nonlinearity, nonlinear dispersion, and absorption of sound in micro-inhomogeneous materials," *J. Acoust. Soc. Am.* **103**, 3216–3226 (1998).
- ¹⁸ V. Gusev, "Parametric acoustic source in a medium with hysteretic quadratic nonlinearity," *Acoust. Lett.* **22**, 30–34 (1998).
- ¹⁹ V. Gusev, "Nonlinear plane waves, induced by moving spatially distributed monochromatic sources," *Vestn. Mosk. Univ., Fiz., Astron.* **36**, 7–12 (1981) [English transl.: *Moscow Univ. Phys. Bull.* **36**, 7–12 (1981)].
- ²⁰ V. E. Gusev and A. A. Karabutov, *Laser Optoacoustics* (AIP, New York, 1993), Chap. 6.

- ²¹V. Gusev, H. Bailliet, P. Lotton, and M. Bruneau, "Interaction of counter-propagating waves in media with nonlinear dissipation and in hysteretic media," *Wave Motion* **29**, 211–221 (1999).
- ²²V. Gusev, "Parametric attenuation and amplification of acoustic signals in the media with hysteretic quadratic nonlinearity," *Phys. Lett. A* **271**, 100–109 (2000).
- ²³V. Gusev, "Quasi-synchronous excitation of acoustic waves in materials with hysteretic quadratic nonlinearity," *Wave Motion* **38**, 241–250 (2003).
- ²⁴I. V. Miloserdova and A. I. Potapov, "Nonlinear standing waves in a rod of finite length," *Akust. Zh.* **29**, 515–520 (1983) [English transl.: *Sov. Phys. Acoust.* **29**, 306–309 (1983)].
- ²⁵M. P. Soerensen, P. L. Christiansen, and P. S. Lomdahl, "Solitary waves on nonlinear elastic rods. I," *J. Acoust. Soc. Am.* **76**, 871–879 (1984).
- ²⁶L. Knopoff, "Attenuation of elastic waves in the Earth," in *Physical Acoustics*, edited by W. P. Mason (Academic, New York, 1965), Vol. III, Pt. B, Chap. 7, pp. 297–325.

On the stability of the effective apodization of the nonlinearly generated second harmonic with respect to range

Russell J. Fedewa,^{a)} Kirk D. Wallace, and Mark R. Holland

Laboratory for Ultrasonics, Department of Physics, Washington University, St. Louis, Missouri 63130

James R. Jago, Gary C. Ng, Brent S. Robinson, and Matthew R. Rielly

Philips Medical Systems, Bothell, Washington 98041

James G. Miller^{b)}

Laboratory for Ultrasonics, Department of Physics, Washington University, St. Louis, Missouri 63130

(Received 27 April 2004; revised 11 January 2005; accepted 29 January 2005)

The concept of an *effective apodization* was introduced to describe the field pattern for the nonlinearly generated second harmonic ($2f$) within the focal zone using a linear propagation model. Our objective in this study was to investigate the validity of the concept of an effective apodization at $2f$ as an approach to approximating the field of the second harmonic over a wide range of depths. Two experimental setups were employed: a vascular imaging array with a water path and an adult cardiac imaging array with an attenuating liver path. In both cases the spatial dependencies of the ultrasonic fields were mapped by scanning a point-like hydrophone within a series of planes orthogonal to the propagation direction. The sampling distances were located before, within, and beyond the focal zone. The signals were Fourier transformed and the complex values at $2f$ were linearly backpropagated to the transmit plane in order to obtain an effective apodization. The measured results demonstrated a relatively constant effective apodization at $2f$ as a function of propagation distance. Finite amplitude computer simulations were found to be in agreement with these measurements. Thus the measure of the effective apodization at $2f$ provides an approximation to the second harmonic field outside the focal zone. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1874612]

PACS numbers: 43.25.-x, 43.25.Jh, 43.25.Cb, 43.60.Fg [MFH]

Pages: 1858–1867

I. INTRODUCTION

Our aim in this paper is to investigate the concept of an effective apodization as a tool for approximating the field pattern of the nonlinearly generated second harmonic ($2f$) portion of a finite amplitude ultrasonic field using a linear propagation model. In previous work the effective apodizations at $2f$ were derived from measurements obtained within the focal zone of the ultrasonic field.^{1–3} In this paper, the determinations of the effective apodizations at $2f$ are extended to include estimates arising from a range of axial positions before, within, and beyond the focal zone. To address the effect of attenuation, two experimental setups using one-dimensional arrays on transmit were studied: a vascular imaging transducer with a lossless water path, and an adult cardiac imaging probe with an attenuating tissue mimicking path. In addition to experimental measurements, results from computer simulations are reported.

The effective apodization at $2f$ is a descriptor of the field pattern of the nonlinearly generated second harmonic. The effective apodization at $2f$ is determined from measurements of the ultrasonic field in a plane perpendicular to the direction of propagation. These measured signals are Fourier transformed and the complex values at the second harmonic frequency are backpropagated to the source plane and the

effective apodization is extracted. (Methods for determining the effective apodization are discussed in greater detail in Sec. II.)

Use of the nonlinearly generated second harmonic has become commonplace in applications of diagnostic imaging.⁴ The shift from using the fundamental to using the second harmonic for image formation is largely driven by observed improvements in image quality. Reasons cited for this improved image quality include the following: reduced sidelobe levels,^{5–7} higher frequencies without the corresponding attenuation penalty,⁷ reduced effects due to phase aberration,^{7–9} and reduced interference from multiple reverberations.^{4,7,8} At this point it is not clear which of the suggested reasons provides the greatest impact on image quality for specific clinical imaging applications.⁴

The nonlinearly generated second harmonic portion of an ultrasonic field has lower sidelobes and a narrower beam than the fundamental portion of the field.^{5,6} The experimental measurements by Ward, Baker, and Humphrey⁶ demonstrated that the full width at half-maximum (FWHM) for the nonlinearly generated harmonics (second, third, and fourth) are related to the fundamental FWHM by $w_n/w_1 = 1/n^m$, where w_n is the width of the harmonic corresponding to the harmonic number n . The value of m was found to be 0.78 for their experimental arrangement. Considering only linear propagation, the value of m would be expected to be unity.¹⁰ Thus the main lobe is slightly wider than would be expected for a simple change in frequency. In addition these authors

^{a)}Current address: Focus Surgery, Inc., 3940 Pendleton Way, Indianapolis, IN 46226.

^{b)}Corresponding author. Electronic-mail: james.g.miller@wustl.edu

noted a decrease in the sidelobe levels for the second harmonic.⁶ In related work by Christopher,⁵ the lateral beam profiles of nonlinearly generated second harmonic fields were compared to the lateral beam profiles of second harmonic fields produced by transmitting at the second harmonic frequency. This work also demonstrated that the nonlinearly generated second harmonic has a wider main lobe and lower sidelobes than a field linearly propagated at the second harmonic frequency.

For a linearly propagating field, the application of a more aggressive transmit apodization results in a wider mainlobe with lower sidelobes.¹⁰ The concept of an effective apodization was introduced as a way of accounting for the field pattern of the nonlinearly generated second harmonic.¹ Changing the transmit apodization has been shown to result in a different effective apodization at the second harmonic frequency² and thus different mainlobe widths and sidelobe levels for the nonlinearly generated second harmonic.¹¹ In addition, the effective apodization at $2f$ has been found to be independent of the F number, focal distance, frequency, and propagation medium for clinically relevant ranges for cases of weakly nonlinear propagation and for data obtained from the focal zone.¹

In a recent publication, the field patterns for the nonlinearly generated second harmonic were compared to linear propagation at $2f$ using the effective apodization at $2f$ determined from the nonlinearly generated second harmonic field.¹² This work demonstrated that the nonlinear field pattern may be approximated by a linear interpretation of the effective apodization at $2f$ descriptor for propagation distances before, within, and beyond the focal zone. These results would indicate that the effective apodization at $2f$ is at most slightly dependent on the propagation distance at which the field is sampled.

For linear propagation, the pulse echo version of the Van Cittert–Zernike theorem relates the transmit apodization to the spatial coherence of backscatter from the focal zone.¹³ By replacing the transmit apodization with the effective apodization at $2f$, the pulse echo version of the Van Cittert–Zernike theorem can be extended to the case of weakly nonlinear propagation (far from shock formation).¹ Thus, the autocorrelation of the effective apodization at $2f$ yields a prediction of the spatial coherence of backscatter that has been shown to be distinct from the spatial coherence of backscatter for the fundamental.^{1–3}

Investigators have reported either an increased^{14,15} or relatively unchanged¹⁶ depth of field for the nonlinearly generated second harmonic component compared with the ultrasonic field produced by transmitting at the same second harmonic frequency (fundamental at second harmonic frequency). The depth of field was determined using three different approaches: single transmit imaging with parallel receive beam forming,^{14,15} Van Cittert–Zernike imaging,¹⁵ and axial pullback measurements.¹⁶ The single transmit imaging with parallel receive beamforming started with a record of signals received on each element.^{14,15} Then offline a receive aperture was synthetically scanned across the received signals to form an image of the transmitted beam. These images demonstrated an extended depth of field for

the second harmonic portion of the field compared with the field produced by transmitting at the second harmonic. Comparable results were achieved by measuring the spatial coherence of backscattered signals using the Van Cittert Zernike theorem to produce an image of the incoherent source.¹⁵ The incoherent source in this case was produced from the interaction of the transmitted signal and the ensemble of randomly distributed, small scatterers. The pullback measurement of received power as a function of propagation distance demonstrated a slight increase in the depth of field for the nonlinearly generated second harmonic field.¹⁶

The apparent disagreement between these reported measurements of the depth of field arises from the different definitions employed by the investigators. The pullback measurement used the positions corresponding to -3 dB with respect to the peak pressure along the propagation axis of a transducer with circular geometry.¹⁶ In contrast, the single transmit imaging with parallel receive beamforming and Van Cittert–Zernike imaging approaches used changes in the beamwidth as a function of axial position.^{14,15} The pullback measurement of depth of field can be affected by attenuation within the sample. In addition, work by others has demonstrated that the pullback definition of the depth of field for the second harmonic can be increased by using multiple foci on transmit.¹⁷ Based on the work in Ref. 17, it appears that having an elevational focus that differs from the azimuthal focus would also extend the pullback measure of the depth of field. However, previous work has shown that the elevational focus has little to no effect on the (azimuthal) effective apodization at $2f$.¹ The pullback measurements supply a measure of the depth of field in terms of the signal to noise ratio. In contrast, the single transmit imaging with parallel receive beamforming and Van Cittert–Zernike imaging approaches provide a depth of field measurement relating to the lateral resolution for the case of having ideal time gain compensation.

In the present work, the concept of an effective apodization is related to the beamwidth as a function of propagation distance and will be shown to agree with the reported observations of an extended depth of field. Both simulations and experimental measurements were performed for the cases of a vascular imaging transducer with a lossless propagation path and a cardiac imaging transducer with an attenuation propagation path. In Sec. II, the methods used for both the experiment and the simulation are discussed in detail. In Sec. III the results are presented in two formats: effective apodizations and apodization widths. In Sec. IV, a discussion of the implications arising from the results is presented. In particular, an effective apodization at $2f$ is used to approximate the field pattern for a nonlinearly generated second harmonic before, within, and beyond the focal zone.

II. METHODS

In order to investigate the field pattern for finite amplitude ultrasonic fields, the effective apodizations at the fundamental ($1f$) and second harmonic ($2f$) were obtained. The effective apodization is determined by measuring the ultrasonic field in a plane perpendicular to the direction of propagation (an orthogonal plane). The complex amplitudes of a

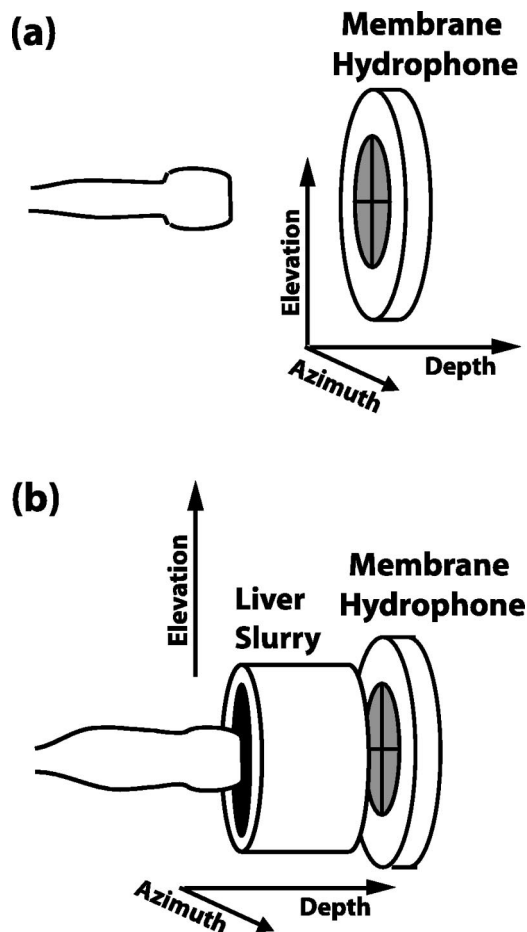


FIG. 1. Experimental arrangements used for measuring ultrasonic fields. Panel (a) depicts the water path setup. Panel (b) depicts the attenuating liver path setup.

single frequency are extracted and used in a linear angular spectrum backpropagation to the source plane.^{18,19} The effective apodization is then obtained as a profile of the magnitude of the backpropagated values. If a frequency within the fundamental band is chosen, the resultant effective apodization at $1f$ is equivalent to the actual transmit apodization.¹⁹ As expected, if a frequency within the nonlinearly generated second harmonic band is chosen, the effective apodization at $2f$ has been shown to be narrower than the transmit apodization for data acquired from the focal zone of the ultrasonic field in both simulation^{2,20,21} and experiment.^{1,2} In the following paragraphs, the experiments and simulations employed to obtain the effective apodizations at $1f$ and $2f$ for propagation depths before, within, and after the focal zone are discussed.

A. Experiment

The two experimental arrangements are shown in Fig. 1. The differences in the two experimental approaches are in the type of transmit array utilized and the type of propagation medium. For the lossless water path depicted in panel (a) of Fig. 1, the ultrasonic field was propagated from a vascular imaging transducer array into water, whereas for the attenuating liver slurry path depicted in panel (b) of Fig. 1, the ultrasonic field was produced by an adult cardiac imaging transducer array. For both setups the ultrasound was propa-

gated perpendicular to the face of a one-dimensional array (i.e., there was no beam steering) using an HDI-5000 ultrasonic imaging system (Philips Medical Systems, Bothell, WA). The transmit F number was approximately 5 for both cases. The two-dimensional pseudoarray scans were performed with a 0.6 mm diameter membrane hydrophone (Sonic Industries, Hatboro, PA) at set distances from the transmit transducer position.

The transmit apodization function was the same for both transducer arrays, and previous work has indicated that the effective apodization at $2f$ at the focus for the adult cardiac imaging transducer agrees with the effective apodization at $2f$ at the focus for the vascular imaging transducer.¹ Thus, the choice of changing both the type of the transducer array and the propagation medium was deemed to be acceptable. The results shown below support this choice.

1. Nonattenuating water path

A drawing of the experimental setup is shown in panel (a) of Fig. 1. Two-dimensional pseudoarray scans were performed at axial positions of 3.0, 22.6, 42.9, 62.6, and 84.3 mm from the face of a one-dimensional vascular imaging array (linear array) transducer. The array elements were focused at 56 mm in the azimuthal direction with a Riesz window apodization, and the elevation direction was focused at 20 mm. Data were acquired by translating the hydrophone receiver in 0.25 mm steps through a 57 by 57 grid (14 mm by 14 mm) and digitizing the RF signal at each position (250 MS/Sec with 8 bits, temporally averaged 32 times). A more detailed description of the experimental setup and data acquisition was published previously.¹

The recorded broadband time-domain signals were Fourier transformed and the complex amplitudes at 5.6 MHz (fundamental) and 11.2 MHz (nonlinearly produced second harmonic) were obtained at each position. These two-dimensional arrays of complex narrow-band values were linearly backpropagated to the plane of the transmit aperture yielding an apparent source from which the effective apodization was determined. (See Refs. 1, 2 for further details on this analysis approach.) The effective apodization for each axial position was determined by averaging the azimuthal profiles contained within the -3 dB elevational source width. Before averaging each of these profiles was normalized by their respective maximum value. The standard deviation of these normalized azimuthal profiles was computed and used as an error measurement.

2. Attenuating liver slurry path

The drawing in panel (b) of Fig. 1 depicts the setup for the one-dimensional adult cardiac imaging transducer array (phased array) with the attenuating liver slurry path. This array was focused at a depth of 100 mm in the azimuthal direction with a Riesz window apodization. The focus in the elevational direction was 47 mm. Two-dimensional pseudoarray scans were performed at axial positions of 2.5, 35, 53, 83, 103, and 162 mm from the face of phased array transducer. A detailed description of the experimental setup, data acquisition, and liver slurry phantom can be found in a

TABLE I. Propagation pathlengths in water and attenuating liver slurry phantom for the through transmission signals associated with the attenuating liver path setup.

Total propagation distance (mm)	Liver slurry phantom length (mm)	Water pathlength (mm)
2.5	...	2.5
35	33.0	2.0
53	50.5	2.5
83	80.0	3.0
103	100.5	2.5
162	159.0	3.0

previous publication² that used the data obtained at the 2.5, 53, and 103 mm axial positions. The propagation distances for the attenuating phantom and nonattenuating water for each axial position from the phased array are listed in Table I.

The effective apodization was obtained using the same approach employed for the nonattenuating water path dataset. In this case, the complex narrow-band data at 1.76 MHz (fundamental) and 3.52 MHz (nonlinearly produced second harmonic) were extracted from the measured rf signals.

B. Simulation

The two experimental transmit configurations (nonattenuating water path and attenuating liver path) were simulated using an approach that incorporated the nonlinear Burgers equation into an angular spectrum propagation model.^{22–24} Each experimental setup was simulated with a Riesz window applied in the azimuthal direction and a rectangular window applied in the elevational direction. The axial propagation step size was 2 mm.

The water path simulation propagated a 51.2 mm by 51.2 mm (256 points by 256 points) source plane a distance of 100 mm. The azimuthal direction was geometrically focused at 56 mm for a speed of sound of 1.54 mm/ μ s corresponding to the programmed speed of sound for the imaging system. We chose acoustic parameters representative of nonattenuating water with the speed of sound set to 1.49 mm/ μ s, the attenuation coefficient set to zero, and the nonlinearity parameter (β) set to 3.6.²⁵ The fundamental frequency was chosen to be 5.6 MHz based on the nonattenuating water path experiment.

The simulation, for an attenuating liver path, propagated a 76.8 mm by 76.8 mm (256 points by 256 points) source plane a distance of 166 mm. The azimuthal direction was geometrically focused at 100 mm, assuming a 1.54 mm/ μ s speed of sound as per the clinical imaging system. We chose an ultrasonic speed of 1.53 mm/ μ s and a slope of attenuation of 0.59 dB/cm/MHz to characterize the attenuating liver slurry phantom within the simulation in agreement with the measured values. A representative value for the nonlinearity parameter (β) of 4.5 was chosen to fall within a range of reported bovine liver measurements.²⁵ The fundamental frequency was chosen to be 1.76 MHz, corresponding to the attenuating liver path experiment.

For both simulations, the complex values for the fundamental and second harmonic frequencies were saved at each 2 mm propagation step. The effective apodization for the fundamental ($1f$) and the second harmonic ($2f$) were determined at each propagation step, as described above. The magnitudes of the effective apodizations were normalized to a peak value of one.

III. RESULTS

Two types of results are presented in this paper: (a) The effective apodizations at $1f$ and $2f$ are compared for a Riesz window transmit apodization, and (b) the widths of the effective apodizations as functions of axial distance.

A. Effective apodizations

The experimentally determined effective apodizations are displayed in Fig. 2 for the water path and for the attenuating liver path. The effective apodizations at $1f$ in panel (a) and panel (c) of Fig. 2 overlie the transmit apodization (Riesz window) for both experimental setups. The effective apodizations at $2f$ in panel (b) of Fig. 2 are all narrower than the Riesz window transmit apodization. Similar results are depicted in panel (d) of Fig. 2 for the attenuating liver path. The lone exception is the measurement at 2.5 mm.

The measurement at 2.5 mm for the attenuating liver path appears to overlie the transmit apodization for positions near the edge of the transmit aperture. There is very little production of second harmonic content arising from nonlinear propagation up to this modest distance of 2.5 mm from the array. A possible explanation for the unexpected second harmonic content is that the ultrasonic source is transmitting a weak (undesirable) signal at the second harmonic frequency. This weak second harmonic signal is attenuated relatively quickly within the attenuating liver slurry phantom. Thus we suggest that the results at 35, 53, 83, 103, and 162 mm are dominated by the nonlinearly produced second harmonic frequency. (This statement is supported by comparisons with the simulation results in the following paragraphs.) The second harmonic results for the water path do not exhibit this second harmonic feature for a field sampled at 3 mm. This suggests that the effect seen in the attenuating liver path may be a result of the transmit characteristics for the adult cardiac imaging array as compared to the vascular imaging array used for the water path.

In Fig. 3, the simulated effective apodizations are shown for the fundamental and the nonlinearly generated second harmonic for both the nonattenuating water path and the attenuating liver path, respectively. The displayed values have been normalized to a peak value of one at each transmit propagation distance. The results for the effective apodization at $1f$ remain unchanged with propagation distance (as expected). Remarkably, the nonlinearly generated second harmonic results are also relatively constant with depth, displaying only a slight decrease in width as the transmit propagation distance increases.

The simulation results (shown in Fig. 3) and experimental results (shown in Fig. 2) for the effective apodizations at $2f$ are compared in Fig. 4. At each distance there is reasonable agreement with the simulation results for both the at-

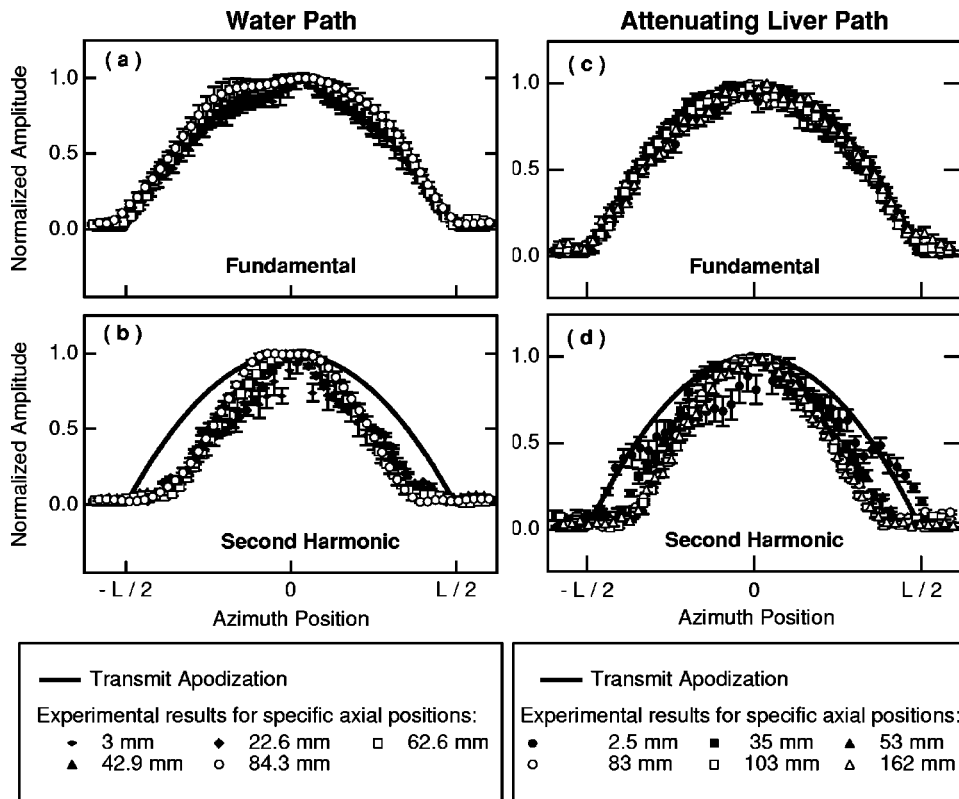


FIG. 2. Experimental results for the effective apodization are displayed for the water path [panels (a) and (b)] and for the attenuating liver path [panels (c) and (d)]. The top row compares the measured effective apodization at $1f$ [5.6 MHz for panel (a) and 1.76 MHz for panel (c)] for each propagation distance. The bottom row compares the measured effective apodizations at $2f$ [11.2 MHz for panel (b) and 3.52 MHz for panel (d)] for all propagation distances from the transmit apodization that were sampled. Error bars represent the standard deviation.

tenuating liver path and the nonattenuating water path. The worst agreement occurs at the shortest propagation distances where the signal to noise ratio of the measured second harmonic is at its lowest.

B. Width of the effective apodization

An estimate of the aperture width for the effective apodizations at the fundamental ($1f$) and second harmonic ($2f$) was determined by finding the positions corresponding to 20% of the peak amplitude in the profiles (approximately a -14 dB level). This metric will be referred to as the *width of the effective apodization*. For a Riesz apodization of full width L , the width of the effective apodization is computed

to be $0.89 L$. For the simulation results, the widths of the effective apodizations were determined from the values displayed in Fig. 3. In order to obtain a measure of the variation in the experimental measurement, a variation of the analysis described in Sec. II was performed. The width was determined for each profile contained within the -3 dB elevational source width before averaging the profiles together. These measurements of the apodization width were then averaged and the standard deviation is reported as the error.

In Fig. 5, the widths of the effective apodizations are displayed as a function of the backpropagation distance for both measured and simulated results. Two normalizations were performed on the data in order to facilitate a comparison

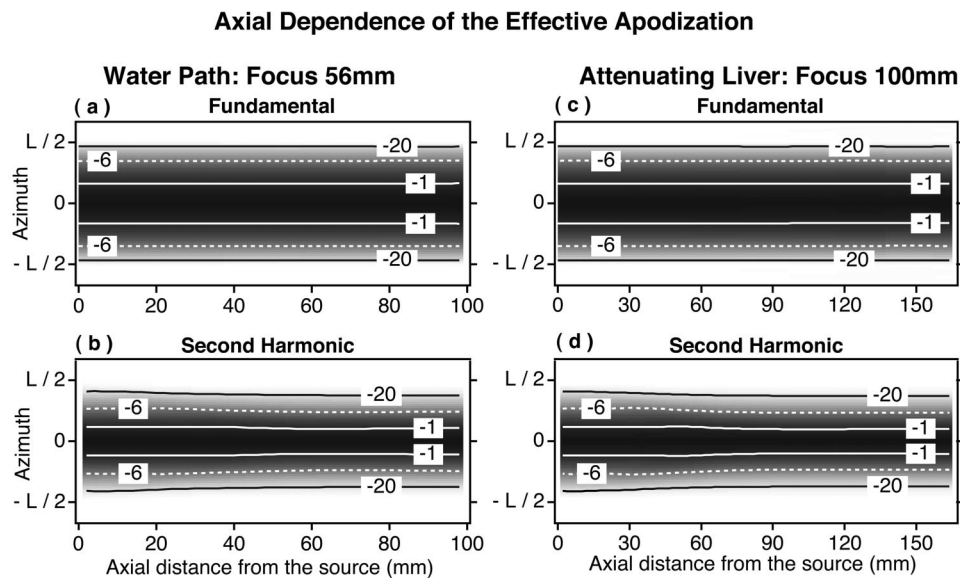


FIG. 3. Results for simulations of the water path and the attenuating liver path are shown. The panels depict the effective apodizations at $1f$ and $2f$ as a function of the backpropagation distance. The top row displays the results for the fundamental. As expected for linear propagation, the contours are horizontal lines. The bottom row displays the results for the nonlinearly generated second harmonic. The effective apodizations for both $1f$ and $2f$ were normalized by their peak value for each backpropagation distance. Contour labels are in units of dB and the linear gray scale is mapped to black for a value of 1.0 and white for a value of 0.0. [The left axis is in units of the full width of the transmit aperture (L).]

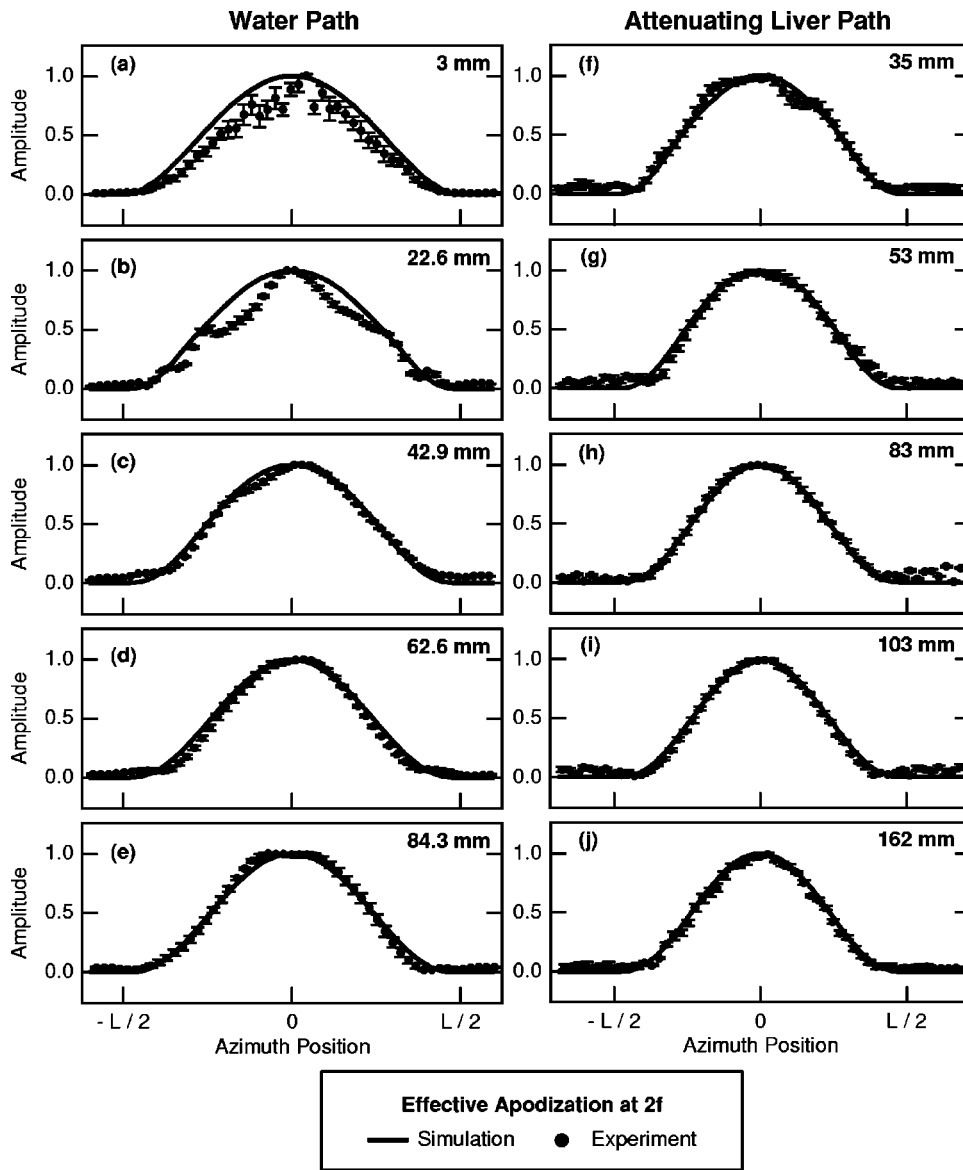


FIG. 4. Effective apodizations at $2f$ for both simulation (solid lines) and experiment (circles) are compared for the water path (first column) and the attenuating liver path (second column). The geometric focus was at 56 mm for the water path and at 100 mm for the attenuating liver path. Error bars represent standard deviation.

son of the results for the two experimental setups. The widths of the effective apodizations were normalized by the transmit aperture full width (L) and the propagation distance was normalized by the geometric focal distance (F).

The widths of the effective apodizations for the fundamental ($1f$) are shown in panel (a) of Fig. 5 and exhibit good agreement between the experiment and simulation for the two different experimental setups. In panel (b) of Fig. 5, the widths of the effective apodizations for the nonlinearly generated second harmonic ($2f$) are displayed. The second harmonic experimental results show reasonable agreement with the simulation results and indicate a slight reduction in width as a function of propagation distance. Comparing panels (a) and (b), the widths for the nonlinearly generated second harmonic are distinctly less than the widths for the fundamental.

The agreement among the widths of the effective apodizations for the second harmonic is notable because (1) a linear array transducer with a geometric focus of 56 mm was compared with a phased array transducer with a geometric focus of 100 mm. (2) The propagation medium for the

water path was nonattenuating, whereas the propagation medium for the attenuating liver path exhibited a 0.59 dB/cm/MHz slope of attenuation. (3) The second harmonic frequencies were 3.52 and 11.2 MHz for the attenuating liver path and water path, respectively.

IV. DISCUSSION

The results from both simulation and experiment demonstrate that the effective apodization of the nonlinearly generated second harmonic field is narrower than the transmit apodization, as can be seen in Fig. 2, Fig. 3, and Fig. 5. The effective apodization at $2f$ exhibits a slight decrease in width as a function of propagation distance, which is much smaller than the difference between any of the effective apodizations at $2f$ and the transmit apodization. In Fig. 5, the most rapid variation in the effective apodization at $2f$ appears to be at a distance around four tenths of the geometric focal distance. Around the focal position and beyond the effective apodization at $2f$ is relatively invariant.

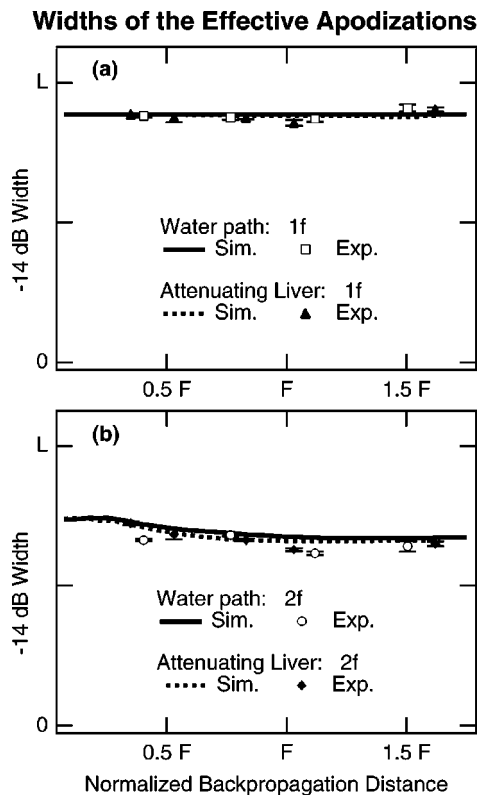


FIG. 5. Widths of the effective apodizations are displayed for the fundamental (a) and the nonlinearly generated second harmonic (b). The widths correspond to the difference in positions determined by the -14 dB value with respect to the peak value for each effective apodization. The propagation distance was normalized by the geometric focus (F). Error bars represent the standard deviation.

The potential influence of the hydrophone's active element size on measurements of the effective apodization can be addressed by examination of Fig. 4 and Fig. 5. Here the results of measurements using a finite-sized aperture hydrophone are compared with the simulation results that are based on point sampling of the simulated field. The agreement between the experimental and simulation results suggest that the finite size of the hydrophone introduced relatively little error. In addition, in the context of this study, the agreement between simulation and measured results shown in Fig. 4 and Fig. 5 suggests that the effective apodization from either pulsed wave (experiment) or continuous wave (simulation) sources exhibit comparable results.

In the following two subsections, the applicability of using the effective apodization at $2f$ to approximate the nonlinearly generated second harmonic is investigated. In the first subsection the field pattern at the second harmonic frequency is compared for two simulations with different transmit conditions: (1) transmitting at the fundamental frequency using a Riesz window apodization; (2) transmitting at the second harmonic frequency using the effective apodization at $2f$ obtained from the results of the first transmit condition. In Sec. IV B the effective apodization at $2f$ is related to reports of an increased depth of field for the nonlinearly generated second harmonic.^{14,16,26} A third subsection contains a discussion about the phase of the effective apodization at $2f$.

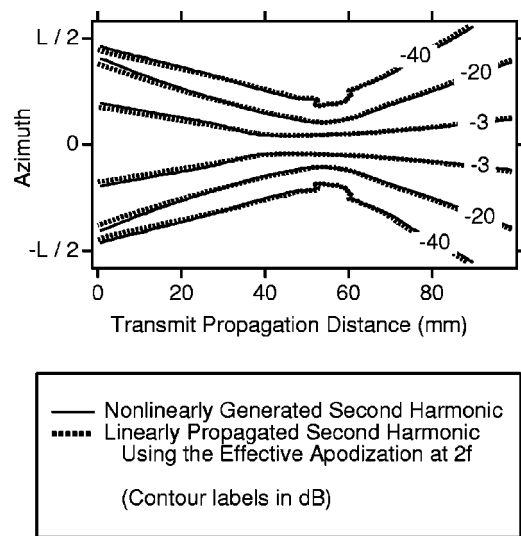


FIG. 6. Contour plots of the normalized field pattern from simulation demonstrating the concept of an effective apodization. The second harmonic portion of an ultrasonic field created by nonlinear propagation of a fundamental transmit with a Riesz window transmit apodization is compared to the ultrasonic field for a linearly propagated signal at the second harmonic frequency that was transmitted using the effective apodization at $2f$ associated with the Riesz window. At each propagation distance the simulation results have been normalized and the contours are reported in units of dB.

A. Estimation of the field pattern

In Fig. 6, simulated field patterns are compared for a nonlinearly generated second harmonic and linear propagation at the second harmonic frequency with the effective apodization determined from the nonlinear field. A simulation of nonlinear propagation in a water path was performed with a fundamental frequency of 5.6 MHz and a Riesz window transmit apodization. The nonlinearly generated second harmonic portion of the field was extracted at every 0.5 mm axial increment. At each propagation distance the power of the second harmonic portion of the field was computed in dB relative to the peak. The complex values of the second harmonic field at the geometric focal position of 56 mm were used to determine the effective apodization at $2f$ (see Sec. II). This effective apodization at $2f$ was then used as a transmit apodization for transmitting at 11.2 MHz. The linear fundamental result from transmitting at 11.2 MHz with the effective apodization at $2f$ overlays the nonlinear result in Fig. 6. Further evidence of this agreement is shown using both experiment and simulation in a recent publication.¹²

The close agreement between the results in Fig. 6 demonstrates the potential usefulness of the concept of an effective apodization at $2f$ for beam formation design. Tool sets developed for linear analysis (such as k -space analysis,^{27,28} spatial impulse response,^{29,30} and the Fraunhofer approximation³¹) may be used to investigate weakly nonlinear propagation by applying the concept of an effective apodization. The main assumption involved in the extension of these tools is that the system has an adequate signal to noise ratio for the second harmonic portion of the field. In addition, the analysis in this paper has assumed no aberration of the signal. The effect of increasing the transmitted pressure on the effective apodization at $2f$ has not been investigated. Thus, the stated restrictions to weakly nonlinear

propagation is a limitation of the current and previous studies.^{1-3,12}

B. Depth of field estimation

The widths of the effective apodizations for the nonlinearly generated second harmonic were found to be considerably less than the transmit aperture widths, as seen in Fig. 5. Reducing the extent of the transmit apodization results in extending the depth of field.¹⁰ Other investigators have observed the apparent extended depth of field for the nonlinearly generated second harmonic portion of the beam as compared with transmitting at the second harmonic frequency using the same transmit apodization.^{14,16,26} Our measurements and simulations of the effective apodization at $2f$ also imply that the depth of field for the nonlinearly generated second harmonic is longer than the depth of field for a transmitted second harmonic that utilized the original transmit apodization.

An estimate of the extension in the depth of field associated with the nonlinearly generated second harmonic is obtained by considering the reduced width of the effective apodization at $2f$. Beginning with a geometric definition of the depth of field (D),

$$D = 7\lambda \left(\frac{F}{L} \right)^2, \quad (1)$$

where F is the geometric focal distance, λ is the wavelength, and L is the aperture width, assuming uniform apodization.¹⁰ Using this equation, the change in the depth of field for transmitting with a Riesz window can be compared to transmitting with the corresponding effective apodization at $2f$. For each case, the same frequency and focal distance are used. Thus, the only difference in the depth field resides in the aperture width (L). The aperture width is computed from the simulation results for the effective apodization at $2f$ for a Riesz transmit apodization. In order to estimate the aperture width, the integral of each apodization is computed and then normalized by a uniform apodization of width L (the transmit aperture width). For the Riesz window and the corresponding effective apodization at $2f$, the resulting values are 0.66 and 0.47, respectively. Thus, the effective apodization at $2f$ indicates an increase in the depth of field of ~ 2.0 times the depth of field obtained by transmitting at $2f$ with the original transmit apodization.

The concept of the effective apodization at $2f$ implies that the depth of field for the nonlinearly generated second harmonic is comparable to the depth of field for the fundamental ($1f$) portion of the field. The difference in frequency is approximately countered by the effective apodization at $2f$ being narrower than the transmit apodization. This estimation for the depth of field does not account for the following: (1) slight changes in the shape of the effective apodization at $2f$ as a function of propagation distance; (2) errors associated with a calculation of the equivalent widths of nonuniform apodization windows; (3) possible changes in the relative phase (geometric focus) of the effective apodization at $2f$ as a function of propagation distance. A brief discussion on the last point is given below.

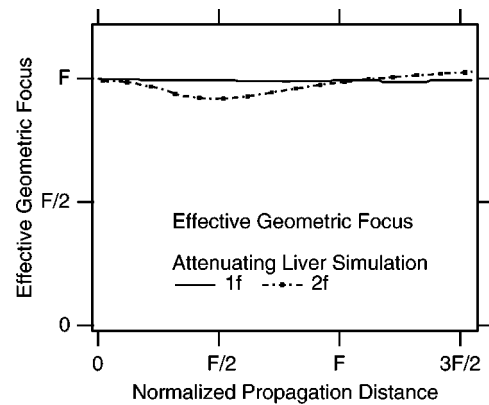


FIG. 7. *Effective geometric focus* as a function of propagation distance for the simulation results of the attenuating liver path are shown.

C. Changes in the phase of the complex effective apodization as a function of propagation distance

The effective apodization is defined above as the magnitude of backpropagated complex values. In this section the phase of these complex values is discussed. Saito reported changes in the phase of the second harmonic with respect to the fundamental as a function of transmit propagation distance for a geometrically focused single element transducer receiving signals from a geometrically focused single element transmitter.³² By analogy, these reported changes may affect the phase of the complex effective apodization at $2f$ as a function of propagation distance. The phase of the complex effective apodization is comparable to the geometric focus for a linearly propagating field. Thus, the nonlinearly generated second harmonic field may be expected to exhibit a dynamically changing *effective geometric focus* as a function of propagation distance.

The phase of the effective apodization can be related to the geometric focus of the apodization. For the simulations described in Sec. II, the *effective geometric focus* was obtained from the phase information within the complex effective apodizations as a function of propagation distance using

$$\Theta(x) = k\sqrt{R^2 + (x-A)^2} - B, \quad (2)$$

where Θ is the phase as a function of aperture position, k is the wave number, R is the geometric focus, and x is the aperture position. The coefficients A and B account for beam steering and a constant phase shift, respectively. A chi-square minimization approach was implemented in the software package IgorPro (WaveMetrics, Inc., Lake Oswego, OR) using Eq. (2) and the simulated complex effective apodization results. The values of k and x were fixed while the values of A , B , and R were determined from the curve fitting routine. The *effective geometric focus* is defined as the resulting value of the geometric focus (R) obtained from the phase of the complex effective apodization.

The *effective geometric focus* predicted from the simulation of the attenuating liver path is shown in Fig. 7. The *effective geometric focus* at $1f$ displays agreement with the known result of the actual geometric transmit focus (F) (100 mm in this case). The *effective geometric focus* at $2f$ demonstrates a difference of, at most, 8% from the value of F and crosses the actual transmit geometric focus at a propagation

distance slightly greater than the actual geometric focus (F). Thus the *effective geometric focus* at $2f$ demonstrates a slight variation as a function of the propagation distance. This variation may result in slightly extending the depth of field for the nonlinearly generated second harmonic. The previous choice to concentrate on the magnitude of the complex effective apodization is supported by the results shown in Fig. 6, where the field pattern of the nonlinearly generated second harmonic is fairly well approximated by a linear propagation using only the magnitude of the complex effective apodization at $2f$.

The effective geometric foci were also determined experimentally. However the errors in the measurements at both $1f$ and $2f$ were much greater than the differences exhibited in Fig. 7 and did not follow any trend as a function of propagation distance. Slight shifts in phase were found to have a significant effect on the value of the *effective geometric focus*. The experimental results were not strong enough to either support or refute the variation in the simulated *effective geometric focus* results shown in Fig. 7.

V. CONCLUSIONS

The concept of effective apodization provides a descriptor of the nonlinearly generated second harmonic portion of an ultrasonic field in terms of linear propagation. The results in Fig. 2 and Fig. 3 show that the widths of the effective apodization at $2f$ are narrower than the effective apodization at $1f$, in agreement with previous measurements.¹⁻³ Furthermore, for an effective apodization at $2f$ obtained from measurements in the focal zone, reasonable agreement to the nonlinearly generated field is observed not only within the focal zone but also well before and beyond the focal zone, as demonstrated in Fig. 6.

In Fig. 5, the widths of the effective apodizations for the nonlinearly generated second harmonic are grouped together for the two different experimental setups for both simulation and experiment. Both experimental setups used a Riesz window transmit apodization and had F numbers of approximately 5. The agreement between the widths is remarkable given that the two experimental setups had different frequencies, different propagation media, different focal distances, and different clinical transducer arrays. Thus, the results shown in Fig. 5 imply that the concept of the effective apodization at $2f$ provides an estimate of the nonlinearly generated second harmonic field for a broad range of experimental conditions.

The impetus for this work was to further the understanding of the underlying principles associated with the improved image quality sometimes observed with harmonic imaging. The concept of an effective apodization provides a tool that may be useful in developing a better understanding of the nonlinearly generated second harmonic field pattern for a homogenous medium. Although imaging in a clinical setting introduces aberrating propagation media, understanding the case for a homogenous propagation medium provides a reference for investigating more complex propagation paths. Furthermore, an application of the concept of an effective apodization may provide a useful tool in the design of ultrasonic imaging systems.

ACKNOWLEDGMENTS

The authors would like to thank Stephanie Posdamer for production of the drawings shown in Fig. 1. This work was supported in part by Philips Medical Systems and NIH R01 HL072761.

- ¹R. J. Fedewa, K. D. Wallace, M. R. Holland, J. R. Jago, G. C. Ng, M. R. Rielly, B. S. Robinson, and J. G. Miller, "Spatial coherence of the nonlinearly generated second harmonic portion of backscatter for a clinical imaging system," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **50**, 1010-1022 (2003).
- ²R. J. Fedewa, K. D. Wallace, M. R. Holland, J. R. Jago, G. C. Ng, M. R. Rielly, B. S. Robinson, and J. G. Miller, "Spatial coherence of backscatter for the nonlinearly produced second harmonic for specific transmit apodizations," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **51**, 576-588 (2004).
- ³R. J. Fedewa, "Nonlinear ultrasonic propagation through aberrating and nonaberrating media," Ph.D. dissertation, Washington University, St. Louis, MO, 2002, available from UMI, 300 North Zeeb Road, PO Box 1346, Ann Arbor, MI 48106-1346 (<http://www.umi.com/hp/Products/Dissertations.html>).
- ⁴F. A. Duck, "Nonlinear acoustics in diagnostic ultrasound," *Ultrasound Med. Biol.* **28**, 1-18 (2002).
- ⁵T. Christopher, "Finite amplitude distortion-based inhomogeneous pulse echo ultrasonic imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **44**, 125-139 (1997).
- ⁶B. Ward, A. C. Baker, and V. F. Humphrey, "Nonlinear propagation applied to the improvement of resolution in diagnostic medical ultrasound," *J. Acoust. Soc. Am.* **101**, 143-154 (1997).
- ⁷V. F. Humphrey, "Nonlinear propagation in ultrasonic fields: measurements, modelling and harmonic imaging," *Ultrasonics* **38**, 267-272 (2000).
- ⁸M. A. Averkiou, "Tissue harmonic imaging," presented at the *2000 IEEE International Ultrasonics Symposium in San Juan, Puerto Rico, 2000*, Vol. 00CH37121, pp. 1563-1572.
- ⁹K. D. Wallace, B. S. Robinson, M. R. Holland, M. R. Rielly, and J. G. Miller, "Experimental comparisons of the impact of abdominal wall aberrators on linear and nonlinear beam patterns," *2004 IEEE International Ultrasonics Symposium*, Montreal, Canada, 2004, Vol. 04CH37553C, pp. 866-869.
- ¹⁰B. D. Steinberg, *Principles of Aperture and Array System Design* (Wiley, New York, 1976), Chap. 3, pp. 40-57.
- ¹¹D. Kourtiche, L. Ait Ali, A. Chitnalah, and M. Nadi, "Theoretical study of harmonic generation sound beams in case of uniform, exponential and cosinusoidal apertures," presented at the *2001 Proceedings of the 23rd Annual EMBS International Conference*, Istanbul, Turkey, 2001, pp. 2395-2398.
- ¹²K. D. Wallace, R. J. Fedewa, M. R. Holland, G. C. Ng, B. S. Robinson, J. R. Jago, M. R. Rielly, and J. G. Miller, "Measurements comparing the linearly propagated field using an effective apodization and the nonlinearly generated second harmonic field," presented at the *2003 IEEE International Ultrasonics Symposium, Honolulu, HI, 2003*, Vol. 03CH37476C, pp. 453-456.
- ¹³R. Mallart and M. Fink, "The Van Cittert-Zernike theorem in pulse echo measurements," *J. Acoust. Soc. Am.* **90**, 2718-2727 (1991).
- ¹⁴D.-L. D. Liu, P. Von Behren, and J. Kim, "Single transmit imaging," presented at the *1999 IEEE Ultrasonics Symposium in Caesars Tahoe, NV, 1999*, Vol. 99CH37027, pp. 1275-1278.
- ¹⁵B. J. Geiman, "Transmit beam estimation in ultrasonic imaging," Ph.D. dissertation, Duke University, Durham, NC, 2000, available from UMI, 300 North Zeeb Road, P.O. Box 1346, Ann Arbor, MI 48106-1346 (<http://www.umi.com/hp/Products/Dissertations.html>).
- ¹⁶E. W. Cherin, J. K. Poulsen, A. F. W. van der Steen, P. Lum, and F. S. Foster, "Experimental characterization of fundamental and second harmonic beams for a high-frequency ultrasound transducer," *Ultrasound Med. Biol.* **28**, 635-646 (2002).
- ¹⁷P. C. Li and C. C. Shen, "Effects of transmit focusing on finite amplitude distortion based second harmonic generation," *Ultrason. Imaging* **21**, 243-258 (1999).
- ¹⁸P. R. Stepanishen and K. C. Benjamin, "Forward and backward projection of acoustic fields using FFT methods," *J. Acoust. Soc. Am.* **71**, 803-812 (1982).

- ¹⁹M. E. Schafer and P. A. Lewin, "Transducer characterization using the angular spectrum method," *J. Acoust. Soc. Am.* **85**, 2202–2214 (1989).
- ²⁰S. Krishnan, J. D. Hamilton, and M. O'Donnell, "Suppression of propagating second harmonic in ultrasound contrast imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**, 704–711 (1998).
- ²¹T. Christopher, "Source prebiasing for improved second harmonic bubble-response imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **46**, 556–563 (1999).
- ²²K. D. Wallace, "Characterization of the nonlinear propagation of diffracting, finite amplitude ultrasonic fields," Ph.D. dissertation, Washington University, St. Louis, MO, 2001, available from UMI, 300 North Zeeb Road, P.O. Box 1346, Ann Arbor, MI 48106-1346 (<http://www.umi.com/hp/Products/Dissertations.html>).
- ²³P. T. Christopher and K. J. Parker, "New approaches to nonlinear diffractive field propagation," *J. Acoust. Soc. Am.* **90**, 488–499 (1991).
- ²⁴C. J. Vecchio, "Finite amplitude acoustic propagation modeling using the extended angular spectrum method," Ph.D. dissertation, Drexel University, Philadelphia, PA, 1992, available from UMI, 300 North Zeeb Road, P.O. Box 1346, Ann Arbor, MI 48106-1346 (<http://www.umi.com/hp/Products/Dissertations.html>).
- ²⁵R. T. Beyer, "The parameter B/A," in *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, San Diego, CA, 1998), Chap. 2, pp. 25–40.
- ²⁶B. J. Geiman, R. C. Gauss, and G. E. Trahey, "In vivo comparison of fundamental and harmonic lateral transmit beam shapes," presented at the 2000 *IEEE Ultrasonics Symposium in San Juan, Puerto Rico*, 2000, Vol. 00CH37121, pp. 1669–1675.
- ²⁷R. M. Lerner and R. C. Waag, "Wave space interpretation of scattered ultrasound," *Ultrasound Med. Biol.* **14**, 97–102 (1988).
- ²⁸W. F. Walker and G. E. Trahey, "The application of k -space in pulse echo ultrasound," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**, 541–558 (1998).
- ²⁹P. R. Stepanishen, "The time-dependent force and radiation impedance on a piston in a rigid infinite planar baffle," *J. Acoust. Soc. Am.* **49**, 841–849 (1971).
- ³⁰J. A. Jensen, "A model for the propagation and scattering of ultrasound in tissue," *J. Acoust. Soc. Am.* **89**, 182–191 (1991).
- ³¹J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, San Francisco, 1968), Chap. 4, pp. 73–75.
- ³²S. Saito, "A property of the nonlinearly generated second-harmonic component of focused ultrasound detected by a concave receiver," *Ultrasonics* **34**, 537–541 (1996).

Acoustic pulse propagation through a fluctuating stably stratified atmospheric boundary layer

Igor Chunchuzov,^{a)} Sergey Kulichkov, Alexander Otrezov, and Vitaly Perepelkin
Obukhov Institute of Atmospheric Physics, 109017 Moscow, 3 Pyzhevskii Per., Russia

(Received 30 June 2004; revised 30 November 2004; accepted 8 January 2005)

Mesoscale wind speed and temperature fluctuations with periods from 1 min to a few hours significantly affect temporal variability and turbulent regime of the stable atmospheric boundary layer (ABL). Their statistical characteristics are still poorly understood, although the knowledge of such statistics is required when modeling sound propagation through the stable ABL. Several field experiments have been conducted to study the influence of mesoscale wind speed fluctuations on acoustic pulse propagation through the stable ABL. Some results of these experiments are described in this paper. A special acoustic source was used to generate acoustic pulses by the detonation of an air–propane mixture with a repetition period 30 s. The mean wind speed and temperature profiles were continuously measured by Doppler sodar and temperature profiler, whereas mesoscale wind fluctuations were measured by anemometers placed on a 56-m mast. From the measurements of the pulse travel time fluctuations at different distances from the source, the statistical characteristics of the mesoscale wind fluctuations, such as frequency spectra, coherences, horizontal phase speeds and scales, have been obtained. Some of the obtained results are interpreted with the use of a recently developed model for the internal wave spectrum in a stably stratified atmosphere. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1862573]

PACS numbers: 43.28.Bj, 43.28.Gq, 43.28.Vd [LCS]

Pages: 1868–1879

I. INTRODUCTION

There are problems of practical importance in atmospheric acoustics which require taking into account the influence of both the mean stratification and the fluctuations of wind speed and temperature on acoustic signal propagation through the atmosphere. These problems include a localization of various acoustic sources in the atmosphere, a prediction of sound levels from pulse and noise sources, and acoustic remote sensing of the atmosphere. To solve them, one needs to parametrize the statistical characteristics of wind speed and temperature fluctuations needed for calculating the amplitude and phase fluctuations of acoustic waves.

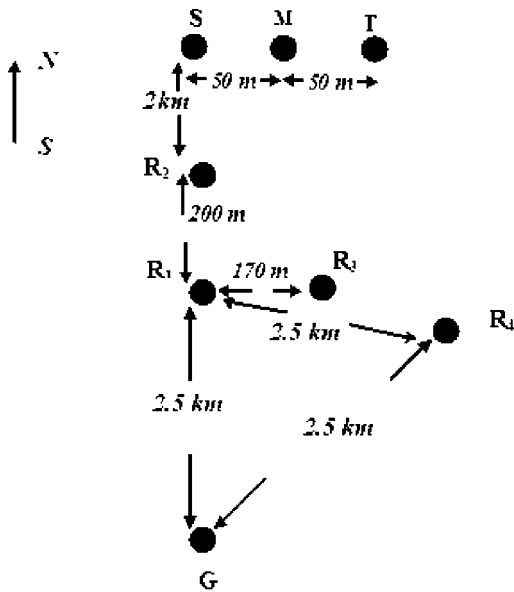
Most sound propagation models use a parametrization of turbulence based on well-known models such as the Kolmogorov–Obukhov model of locally isotropic and homogeneous turbulence (Tatarskii, 1971), the von Karman model (Ostashev *et al.*, 1998), and Gaussian models for both isotropic (Chernov, 1960) and anisotropic turbulence (Wilson and Thomson, 1994).

However, in stably a stratified atmospheric boundary layer (ABL), the parametrization of the turbulence statistics is still a problem, since these characteristics at high Richardson numbers, Ri , cannot be described by means of the existing similarity theories, most of which are based on the assumption that the turbulence is locally homogeneous and stationary (statistically). The widely used models of the ABL based on a large-eddy simulation do not capture flows with very stable stratification (Derbyshire, 1999; Wilson *et al.*, 2001) and fail to simulate layers with the sharp vertical gradients that are often observed in the ABL.

It is recognized now that different types of instabilities of internal gravity waves (IGWs) may be a source of small-scale turbulence and meso-scale eddy structures (such as “cat eyes,” “filaments,” “banks,” and others) in stable ABLs (Chimonas, 1978; Gossard *et al.*, 1985; Danilov and Chunchuzov, 1992; Chunchuzov, 1996; Chimonas, 1999; Galaktionov and Maderich, 1999; Blumen *et al.*, 2001; Anderson, 2002). Such turbulence coexists and continuously interacts with the IGWs, so the statistical properties of these waves affect both the statistics of turbulence and the fine vertical structure of the wind field in a stably stratified ABL (Weinstock, 1984; Einaudi and Finnigan, 1993).

The IGWs themselves induce meso-scale wind speed and temperature fluctuations in the stable ABL with periods from a few minutes to a few hours, and with horizontal scales from hundreds of meters to a few kilometers. These fluctuations cause the corresponding fluctuations of the parameters of acoustic signals propagating through the ABL, such as travel time, amplitude, duration, and the angles of arrival of the acoustic signals. To understand how IGWs influence the statistical characteristics of the parameters of acoustic signals, and the turbulent fluxes of heat and momentum in the stable ABL, one needs to know the form of the frequency-wavenumber spectrum of the wind speed and temperature fluctuations induced by IGWs. Such a need motivated our study of the effects of meso-scale wind speed and temperature fluctuations associated with IGWs on the statistical characteristics of the parameters of acoustic pulses in the stable ABL. The experimental results of this study are described in Sec. II. In Sec. III A these results are interpreted by using a theoretical model for the frequency spectra of the wind speed fluctuations induced by internal waves. Based on this model a theoretical form for the frequency spectrum of

^{a)}Address for correspondence: 2 Putman Ave, Ottawa, ON K1M 1Y9, Canada. Electronic mail: tigra@ifaran.ru and oxana@achilles.net



G – GENERATOR, R_i – RECEIVERS,
S – SODAR, M – MAST, T – PROFILER

FIG. 1. Scheme of the field experiment on acoustic pulse sounding of stable ABL, 9–13 August 2002.

travel time fluctuations is obtained in Sec. III B and compared with the measured spectral forms (Sec. III C). The obtained results of the paper are discussed in Sec. 4.

II. EXPERIMENTAL STUDY OF THE INFLUENCE OF MESO-SCALE WIND SPEED FLUCTUATIONS ON THE PARAMETERS OF ACOUSTIC SIGNALS

A. Acoustic pulse sounding method

To measure meso-scale wind speed and temperature fluctuations, we used an acoustic pulse sounding method previously developed for this purpose (Chunchuzov *et al.*, 1990, 1996, 1997; Chunchuzov, 2002). This method is a type of acoustic tomography of the stable ABL based on the existence of an acoustic wave guide near ground surface due to formation of the temperature inversion and vertical wind shear near the ground surface.

The scheme of the field experiment conducted in August 2002 at the base of the Institute of Atmospheric Physics near Zvenigorod is shown in Fig. 1. It consists of an acoustic pulse generator (G), receivers (R1–R4), Doppler sodar (S), a temperature profiler (T) and a meteorological mast (M) with the acoustic anemometers mounted at heights of 6 and 56 m above ground surface.

Acoustic pulses of stable form were generated by detonation of an air-propane mixture in the special tube shown in Fig. 2. We chose a repetition period of 30 s for sounding of the meso-scale wind speed and temperature fluctuations with temporal scales ranging from 1 min to 1 h. One of the acoustic pulses received at a distance of 20 m from the generator is shown in the upper panel (a) of Fig. 3. When propagating in the acoustic wave guide the initial signal “splits” at a distance of 2.5 km into a set of arrivals A, B, C, and D, as

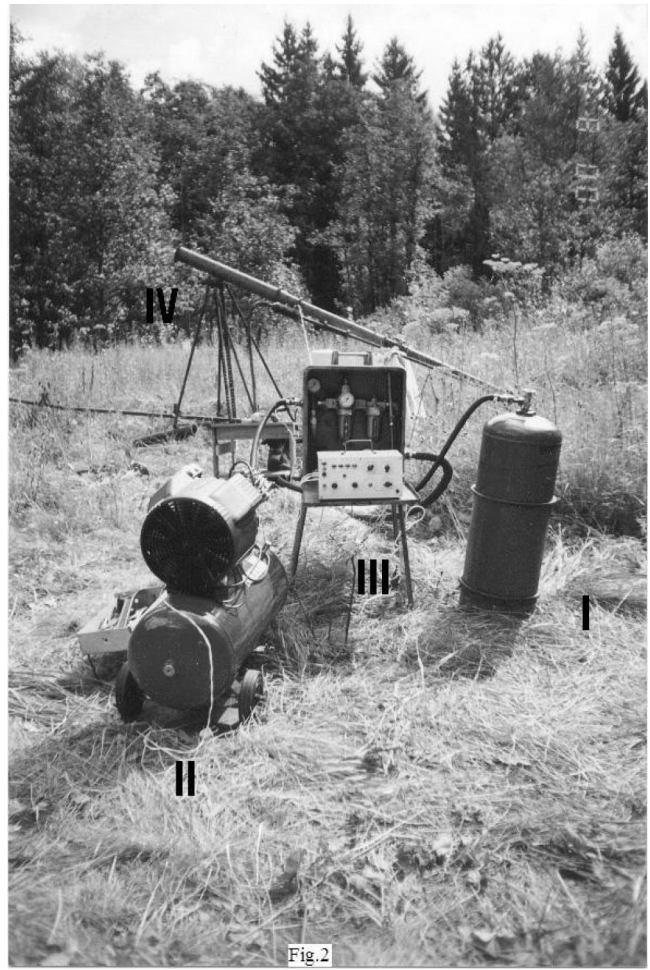


FIG. 2. Detonation generator.

shown in Fig. 3(c). These arrivals are readily identified in the records of the signal [shown in Figs. 3(b)–3(d)] detected by the receivers R1, R2, and R3, which form a 5-m horizontal triangle. Such evolution of the pulse waveform with distance was earlier explained theoretically by using exact solutions of the Helmholtz equation for the sound field in a moving stratified layer (Chunchuzov *et al.*, 1990, 1996).

For each signal radiated at the time t we measured the time interval $\Delta T(t)$ between arrivals (D) and (A) with the accuracy of 2 ms. By measuring temporal variations $\delta T(t) = \Delta T(t) - \Delta T_0$ of the interval $\Delta T(t)$ relative to its value ΔT_0 at some initial moment t_0 , we were able to calculate the temporal variations $\delta c_{\text{eff}}(t) = c_{\text{eff}}(t) - c_{\text{eff}}(t_0)$ of the effective sound speed $c_{\text{eff}} = c + V_e$ averaged over the selected ray trajectory Γ_A connecting the source and each receiver. Here c is the sound speed, V_e is the projection of the wind velocity along the radius-vector from the source to the receiver, and the trajectory Γ_A is thought to have a maximum elevation angle relative to the ground surface and one turning point.

As earlier shown by Chunchuzov *et al.* (1997) the variations $\delta T(t)$ are caused mostly by the variations of the pulse travel time along the ray Γ_A , because their contributions to the value of $\delta T(t)$ are much greater than those from the travel time variations along the nearly horizontal ray path Γ_D for an arrival D. The latter has an elevation angle close to 0.

For small relative variations of the effective sound

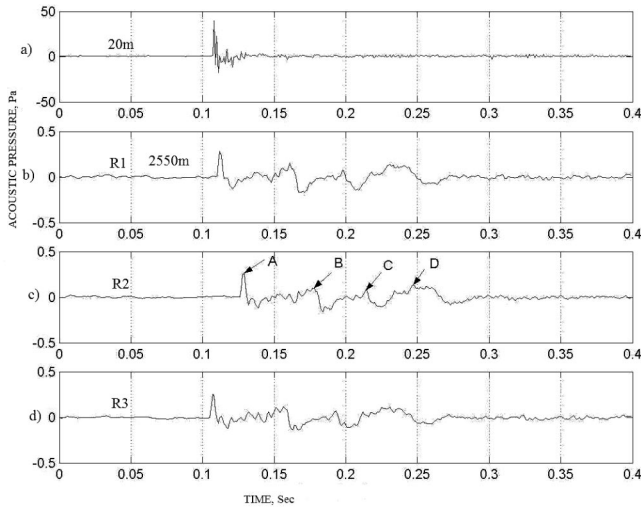


FIG. 3. One of the acoustic signals generated by a detonation source on 10 August 2002, 22:40. This signal at a distance of 20 m is shown in the upper panel (a). The same signal was also recorded far from the source (2550 m for the receiver R1) by a 5-m triangle array R1, R2, and R3. These recordings at the points R1, R2, and R3 are shown in the panels (b), (c), and (d), respectively. A, B, C, and D are different arrivals of the same signal.

speed, $|\delta c_{\text{eff}}(t)/c_{\text{eff}}|$, we can express $\delta T(t)$ as $\int_{\Gamma_A} ds (\delta c_{\text{eff}}(t) \cdot c_{\text{eff}}^{-2}) = \langle \delta c_{\text{eff}}(t) \rangle \int_{\Gamma_A} ds (c_{\text{eff}}^{-2})$, where the second-order terms $O(|\delta c_{\text{eff}}(t)/c_{\text{eff}}|^2)$ are neglected, and $\langle \delta c_{\text{eff}}(t) \rangle$ is the value of $\delta c_{\text{eff}}(t)$ averaged over the initial ray path Γ_A for the arrival A (Chunchuzov *et al.*, 1997). If the relative increase of the effective sound speed from the ground to the height of the ray turning point is small, then in the first-order approximation the initial time interval ΔT_0 may be expressed as $\int_{\Gamma_A} ds ((c_{\text{eff}}(s) - c_0) c_{\text{eff}}^{-2}) = \langle \Delta c_{\text{eff}} \rangle \int_{\Gamma_A} ds (c_{\text{eff}}^{-2})$, where c_0 is the initial value of sound speed c near ground surface $z=0$, and $\langle \Delta c_{\text{eff}} \rangle$ is the average value of the difference $\Delta c_{\text{eff}}(s) = (c_{\text{eff}}(s) - c_0)$ over the ray path Γ_A at the initial moment t_0 (note that under this approximation we have replaced the integration over Γ_D by the integration over Γ_A). As a result the relation between averaged time fluctuations $\langle (\delta c_{\text{eff}})_m(t) \rangle$ and the fluctuations $\delta T(t)$ becomes

$$\langle (\delta c_{\text{eff}})_m(t) \rangle \approx \langle \Delta c_{\text{eff}} \rangle \delta T(t) / \Delta T_0. \quad (1)$$

The value of $\langle \Delta c_{\text{eff}} \rangle$ may be readily estimated for the initial ray trajectories calculated for the wind speed and temperature profiles measured by the sodar and temperature profiler, respectively, so from measured values of $\delta T(t)$ and

ΔT_0 one can calculate from (1) the averaged effective sound speed fluctuations $\langle (\delta c_{\text{eff}})_m(t) \rangle$.

B. Experimental results

For the signal in Fig. 3(c) the acoustic pressure field calculated by a parabolic equation method (Aivilov, 1995) is displayed in Fig. 4. A constructive interference of normal modes forms the ray paths, which are clearly seen in this figure. For the totally reflecting ray (A), with a maximum elevation angle relative to the ground, the height of the turning point is about 99 m. For this height the estimated average increase in the effective sound speed $\langle \Delta c_{\text{eff}} \rangle$ is about 3 m/s.

Since the initial time interval ΔT_0 between arrivals A and D was 70 ms, then from the fluctuations $\delta T(t)$ measured by means of a horizontal triangle array R1–R3, one can calculate the average fluctuations of the effective sound speed $\langle (\delta c_{\text{eff}})_m(t) \rangle$. As shown by Chunchuzov *et al.* (1997), the main contribution to the value of $\langle (\delta c_{\text{eff}})_m(t) \rangle$ comes from the small portion of the ray trajectory Γ_A near its turning point. The time series of the fluctuations $\langle (\delta c_{\text{eff}})_m(t) \rangle$ for the ray paths Γ_A that connect the source G with each of the receivers R1, R2, and R3 are shown in Fig. 5(a). For the same time period Fig. 5(b) shows the time series for the fluctuations of the horizontal wind speed projection (v_x) along the x -axis directed from R1 to R2 (i.e., northward). These are measured by the anemometer at a height of 56 m (the receivers R1–R3 and the mast with the anemometers are shown in Fig. 1).

The obtained time series of $\langle (\delta c_{\text{eff}}) \rangle$ and v_x reveal similar oscillations with the periods ranging from 8 min to about 20 min. To estimate the horizontal speed and direction of the propagation of such oscillations, the receiver R2 was placed 200 m northward (x -axis) from R1, whereas the receiver R3 was placed 170 m eastward from R1 so that these receivers formed a horizontal triangle array (Fig. 1).

The fluctuations of the pulse travel time measured by the triangle array R1–R3 in the night time hours of 13 August 2002 are shown in Fig. 6. For these time series we have calculated the coherences K_{ij} and the corresponding phase spectra $\varphi_i - \varphi_j$ between the points i and j of the triangle, where $i, j = 1, 2, 3$. The coherences K_{ij} and the multi-coherence function defined by Grachev *et al.* (1978) as

$$K_0 = (K_{12}^2 + K_{23}^2 + K_{31}^2 - 2K_{12}K_{23}K_{31} \cos \Sigma \varphi)^{1/2}, \quad (2)$$

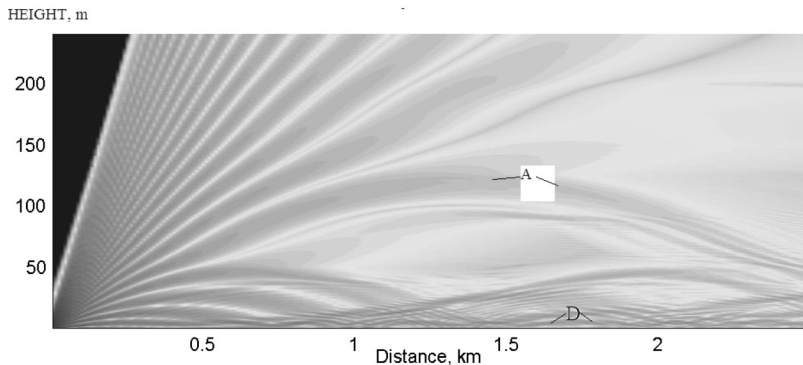


FIG. 4. Ray trajectories for the signal in Fig. 3(b). The height of the turning point for the totally reflected ray with a maximal elevation angle is 99 m (this ray corresponds to the arrival A), the average effective sound speed increase $\langle \Delta c_{\text{eff}} \rangle$ from the ground to this height is 3 m/s, and the time interval between arrivals A and D is 70 ms.

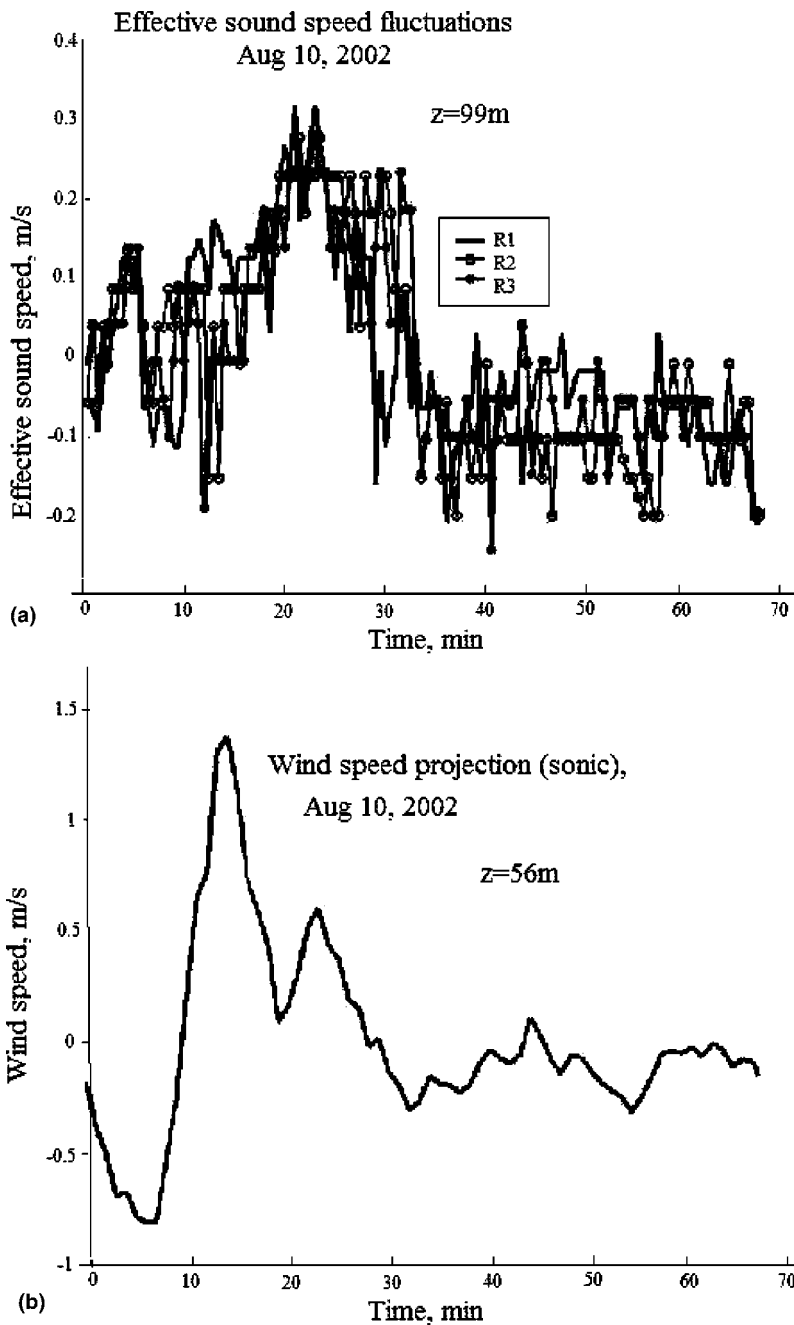


FIG. 5. The time series (a) of the effective sound speed fluctuations $\langle(\delta c_{\text{eff}})_m(t)\rangle$ reconstructed from the travel time measurements by a 5-m triangle array, and the horizontal wind speed fluctuations (b) measured by acoustic anemometer at a height of 56 m.

where $\Sigma\varphi = (\varphi_1 - \varphi_2) + (\varphi_2 - \varphi_3) + (\varphi_3 - \varphi_1)$ is the sum of the phase differences, were estimated by means of standard procedures (Bendat and Piersol, 1967). If $K_{12} = K_{23} = K_{31} = 1$ and $\Sigma\varphi = 0$, then $K_0 = 1$, so the existence within some frequency interval of the maximum of the multi-coherence function K_0 along with the condition $\Sigma\varphi = 0$ may be considered as an indication of the occurrence of wavelike fluctuations within the same interval.

As seen from Fig. 7, all the calculated coherences show a low-frequency peak with the maximum between 0.7 and 0.98 and within a low-frequency interval of 0.6–1.7 mHz, where 1 mHz = 10^{-3} Hz (the corresponding periods are between 10 and 28 min). Within the same frequency range the sum of phase differences $\Sigma\varphi$ becomes close to zero (Fig. 7), and this fact indicates a wavelike character of the observed low-frequency fluctuations. Note that the central frequencies

of the peaks of different coherences K_{ij} may be slightly shifted relative to each other, and this seems to be associated with the azimuthal anisotropy of the speed of the wavelike fluctuations with respect to the mean wind velocity. In this case only those frequencies have been selected, for which all the coherences reached a maximum along with the condition $\Sigma\varphi \approx 0$. Beside the low-frequency peak in the interval 0.6–1.7 mHz, this condition was also met for the frequencies 6.7 mHz (period 2.5 min) and 10 mHz (period 1.5 min). Thus there exists a number of dominant frequencies in the observed fluctuations of the pulse travel time. For these frequencies we have calculated the phase differences between the three turning points of the refracting rays Γ_A connecting a source with the receivers R1, R2, and R3 (the distances between these turning points are $\Delta x = 100$ m and $\Delta y = 85$ m).

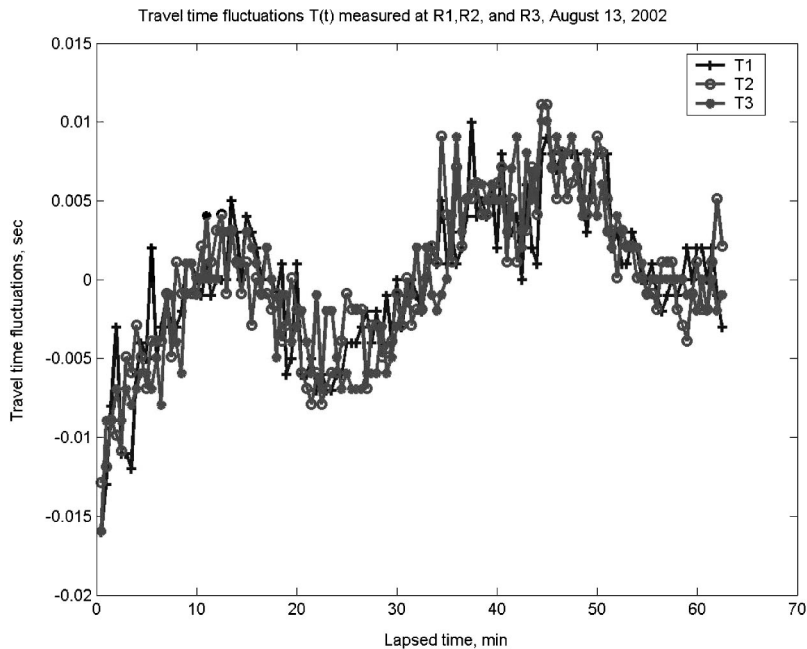


FIG. 6. Pulse travel time fluctuations $\delta T(t)$ measured on 13 August 2002 (22:40–22:50) by a 200-m triangle array R1–R3 (see Fig. 1).

The calculated phase velocities V_p of the effective sound speed fluctuations versus frequency are shown in Fig. 8(a). The value of V_p as seen decreases from 5.5 m/s at a frequency $f=0.6$ mHz (corresponding period is 28 min) to about 2 m/s at $f=2$ mHz (period is about 8 min). The obtained periods and corresponding phase velocities are typical for the internal waves often observed in the stably stratified ABL (Kjelaas *et al.*, 1974; Gossard and Hooke, 1975; Aggarwal *et al.*, 1980). The calculated azimuthal direction of propagation of the wavelike fluctuations is within the interval 100° – 105° with respect to the x axis as shown in Fig. 8(b). The corresponding horizontal wavelengths of these fluctuations $\lambda = V_p/f$ decrease from 8.6 km at a period of 28 min to 960 m at a period of 8 min. At a shorter period of 2.5 min, the azimuth of the fluctuations is about 100° and $\lambda \approx 270$ m.

C. Stratification of the ABL and throposphere during observations

The observed fluctuations within the range of periods (2.5–10 min) may be induced by trapped internal wave modes propagating in the wave guide formed near ground surface by the mean stratification of the Brunt-Väisälä (BV) frequency $N(z)$ (Gossard and Hooke, 1975) and wind speed in the stable ABL. This is seen from the vertical profiles of temperature [Fig. 9(a)], BV frequency $N(z)$ [Fig. 9(b)], and horizontal wind speed [Figs. 9(c) and 9(d)] in the lower 400-m atmospheric layer. Indeed, the value of N^2 decreases from its maximum $N_{\max}^2 \sim 1.2 \times 10^{-3} \text{ rad}^2/\text{s}^2$ at the height of 75 m to its minimum $N_{\min}^2 \sim (0.2\text{--}0.3) \times 10^{-3} \text{ rad}^2/\text{s}^2$, which is between the heights of 350 and 500 m.

Thus, the lower 350-m layer of the ABL may serve as a wave guide for internal gravity wave modes with the frequencies $\omega' = \omega - \mathbf{k}\mathbf{V}(z) > 0$, satisfying the condition $N_{\min} < \omega' < N_{\max}$, where an intrinsic frequency ω' is taken relative to the mean wind speed $\mathbf{V}(z)$ at a given height z , and \mathbf{k} is the horizontal wave number of a trapped mode (Lighthill, 1978).

The corresponding interval of periods $2\pi/\omega'$ for the trapped modes is between the values $2\pi/N_{\max} \approx 3$ min and $2\pi/N_{\min} \approx 6\text{--}8$ min.

Note that above the 70-m layer of strong temperature inversion the value of the mean vertical wind shear is less than 0.02 Hz, and the corresponding Richardson number Ri is between 1 and 3 (this estimate was done for the 50-m vertical resolution of temperature and wind speed). Hence the atmospheric layer between 70 and 350 m was stable both statically and dynamically. The linear frequencies $f = \omega/2\pi$ of the trapped wave modes relative to the ground are in the interval

$$(N_{\min} + \mathbf{k}\mathbf{V})/2\pi < f < (N_{\max} + \mathbf{k}\mathbf{V})/2\pi. \quad (3)$$

It follows from (3) that the wave number values $k = 2\pi f/V_p$ of the observed wavelike fluctuations vary from 7.4 rad/m at $f=0.6$ mHz to 3.1×10^{-3} rad/m at $f=2$ mHz, whereas the direction of the horizontal wind speed \mathbf{V} (relative to the x axis) is between 60° and 170° . In this case the angle α between vectors \mathbf{k} and \mathbf{V} is within the interval 120° – 160° , hence the observed wavelike fluctuations propagate almost opposite to the mean wind velocity.

For $k = 3.1 \times 10^{-3} \text{ m}^{-1}$, $\alpha = 100^\circ$ and for the maximum of wind speed $V = 6$ m/s within a 350-m layer the estimated frequencies for the trapped wave modes are in the range $1.5 \times 10^{-3} \text{ Hz} < f < 4 \times 10^{-3} \text{ Hz}$ (the corresponding periods are between 4 and 11 min). The modes with periods of more than 11 min are not trapped within the lower 350-m layer, so they can freely propagate through the tropospheric layers that are aloft. One possible source of these waves may be an unstable wind shear which occurred in the lower troposphere during the night hours of 13 August 2002. As follows from the radiosonde data obtained during the pulse soundings, an unstable wind shear layer with $Ri < 0.25$ existed in the height range between 1 and 1.8 km. The wind speed in this layer reached a maximum of about 6 m/s, which was close to the horizontal phase speed of the observed wavelike fluctuations

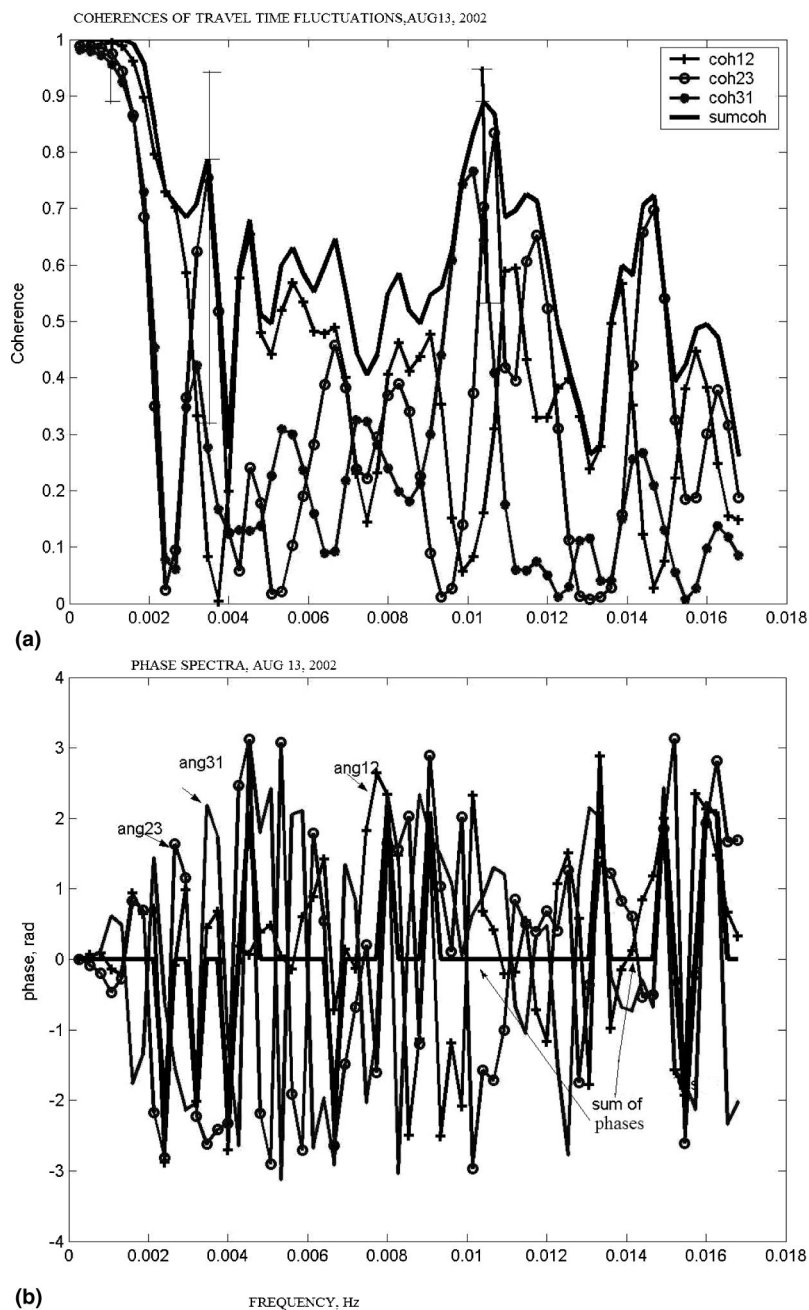


FIG. 7. Coherences and phase spectra for the time series in Fig. 6, 13 August 2002. (a) Coherences between the points R_i and R_j , ($i, j = 1, 2, 3$) and a multi-coherence function (continuous line). Vertical bars show 90% confidence intervals. (b) Phase spectra $\varphi_i - \varphi_j$ and their sum (continuous line).

with a period of 28 min. Such conditions were favorable for the generation of internal waves by an unstable tropospheric wind shear, and for the downward propagation of these waves through a stable ABL to the ground surface (Gossard and Hooke, 1975).

Thus, the frequency range 0.6–1.7 mHz corresponding to the observed peak of coherence may be caused by the untrapped IGWs with periods of 11–28 min propagating downward from tropospheric source. The fluctuations with higher frequencies (periods from 4 to 11 min) may be induced by the trapped wave modes in the lower 350-m layer.

Under the weak mean wind shear the trapped (standing) wave modes may induce wind shears that reach critical values at certain heights (depending on the forms of wind speed and BV-frequency profiles), where the Richardson number Ri becomes less than $\frac{1}{4}$ (Chimonas, 1978; Gossard *et al.*,

1985; Danilov and Chunchuzov, 1992). At these heights, the field of internal wave modes breaks down and generates thin layers of turbulence. During the night time measurements of 13 August 2002, the existence of the trapped IGWs under the weak mean wind shear ($\ll 0.02 \text{ s}^{-1}$) was a likely source of the horizontal layers of intense turbulence in the stable ABL. These layers are seen in the sodar echogram (Fig. 10) during the time period 22:40–00:30 LT as the dark isolated layers at the altitudes of 100–150 m and 300–350 m. Note that these layers are undulating with periods (4–5 min) inherent in the trapped internal wave modes. The vertical displacements of the undulating layers reach values of about 100 m, which are comparable to the vertical length scales of the trapped wave modes. The wave modes with such large relative amplitudes should be strongly nonlinear and, therefore, strongly interact between each other. The importance of nonlinear effects for

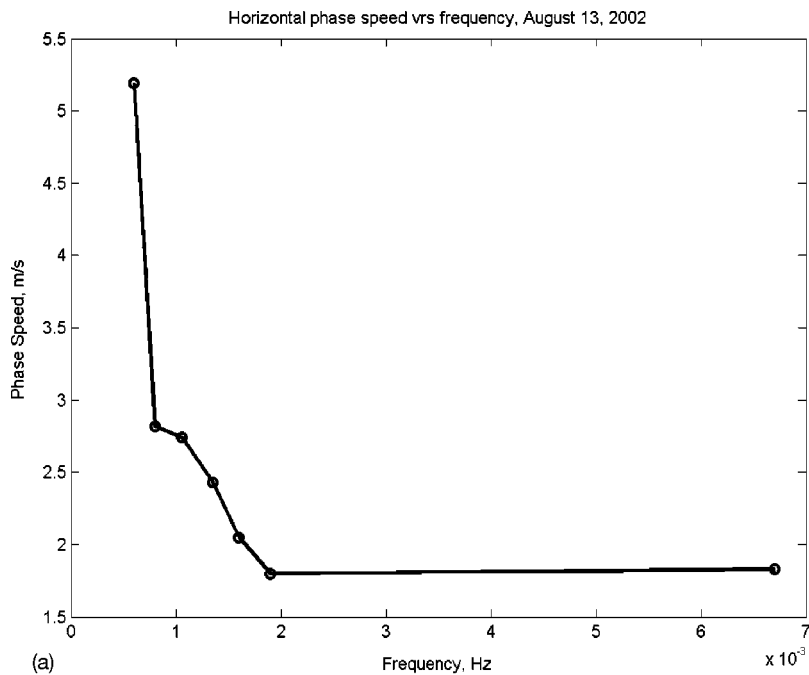
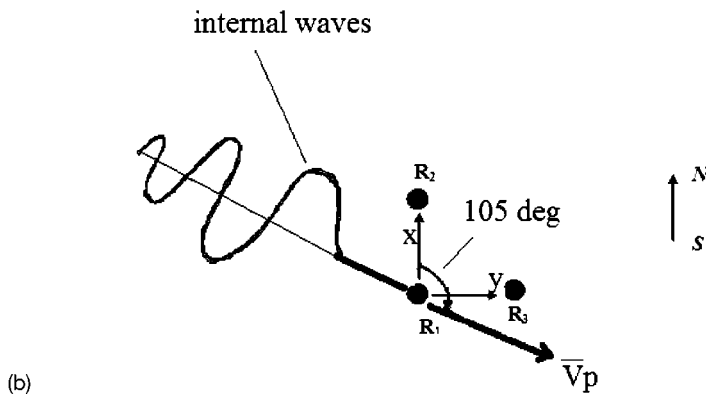


FIG. 8. Horizontal phase speed V_p of the effective sound speed fluctuations (a) (obtained from the time series in Fig. 6) as a function of frequency. (b) Illustration of the propagation of the observed internal waves with respect to the triangle array R1, R2, and R3.



the observed internal waves in the ABL also follows from the comparison (see Sec. III A) of the obtained values of phase speeds for these waves with the rms values of the horizontal wind speed fluctuations.

III. FREQUENCY SPECTRA OF TRAVEL TIME FLUCTUATIONS

A. Model of the temporal internal wave spectrum

The rms values σ of the wind speed fluctuations obtained from the anemometer measurements at a 56-m height during periods of acoustic pulse sounding of stable ABIs (from 10 to 13 August 2002) ranged from 0.1 to 0.5 m/s. At the same time the horizontal phase speeds of the internal waves detected in the stable ABL are about several meters per second. These values are close to the mean horizontal wind speed values in the ABL, so the wave phase velocities relative to the mean wind speed may be comparable to the rms values σ of the wind speed fluctuations. In this case, the advective term $(\mathbf{v}\nabla)\mathbf{v}$ in the Eulerian motion equation becomes comparable in magnitude to the linear term $\partial\mathbf{v}/\partial t$ that describes the acceleration of fluid parcels at a fixed point.

This means that the effects of advection of the IGWs by a fluctuating wind play a key role in the formation of the wave energy distribution among IGWs of different scales in the Eulerian frame of reference. There is also a dynamic nonlinearity of the equation of state that affects the energy spectrum of IGWs along with advective nonlinearity.

The influence of both advective and dynamic nonlinearities of the Eulerian motion equations on shaping of the energy spectrum of IGWs was taken into account in the model of the IGW spectrum in the ocean developed by Allen and Joseph (1989). This model was then adopted and modified for the atmosphere (Churchuzov, 1996, 2002; Hines, 2001). The theory developed in these works predicts universal forms for both spatial and temporal spectra of the wind speed and temperature fluctuations induced by an ensemble of gravity wave modes with randomly independent phases and amplitudes. Some theoretical results on the frequency spectrum of IGWs will be used here to interpret the obtained experimental frequency spectra of the acoustic pulse travel time fluctuations in a stably stratified ABL.

The simplest two-layer model of the stably stratified and nonrotating atmosphere [$N(z)=N_1$ for $0\leq z\leq h$, and $N(z)$

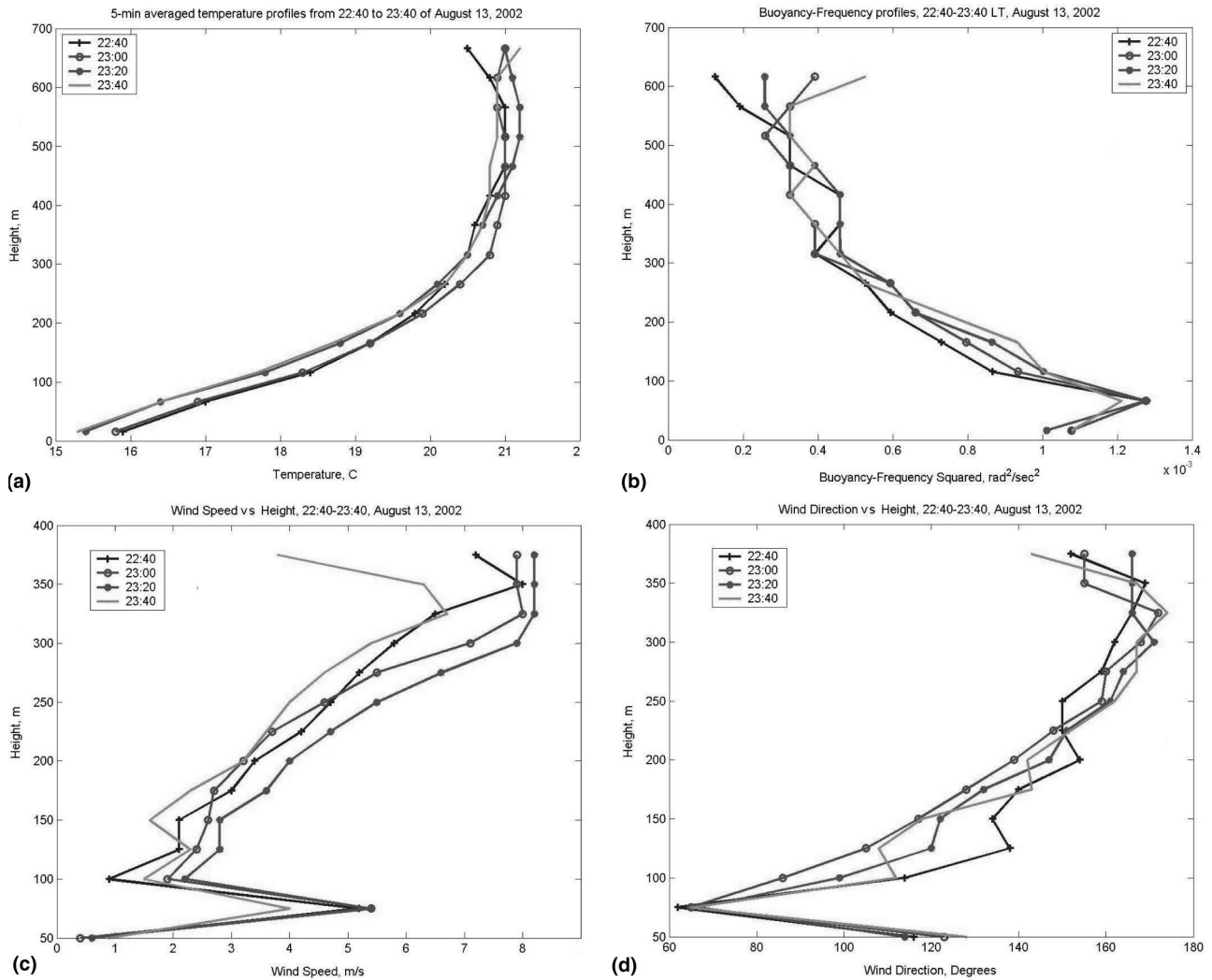


FIG. 9. Vertical profiles of temperature measured by temperature profiler, BV frequency, wind speed, and wind direction (measured by Doppler sodar), 13 August 2002, 22:40–23:50. These profiles are averaged over 5 min. (a) Temperature; (b) BV frequency; (c) wind speed; and (d) wind direction.

$=N_2$ for $h \leq z \leq \infty$, where $N_2 \ll N_1$] was chosen to find an equilibrium energy spectrum of the random internal wave field in the lower atmospheric layer. This model may be used when the lower atmospheric layer with a depth h of about 1–2 km has a mean static stability (over the entire thickness of the layer) N_1 higher than the tropospheric layer aloft (Gossard and Hooke, 1975). This is the case shown in Fig. 11. As seen from the tropospheric BV-frequency profile (Fig. 11) obtained from radiosonde data of 13 August 2002, the value of $N(z)$ decreased with a height z from about 3 mHz

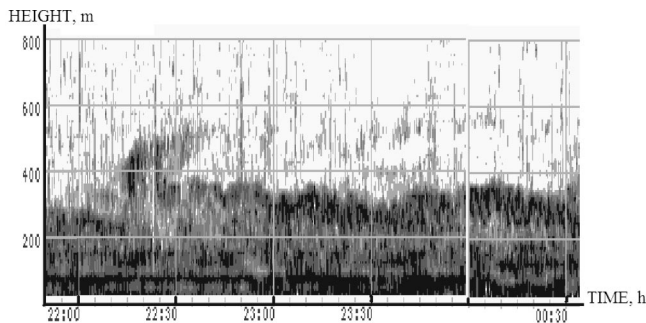


FIG. 10. Sodar echogram during the time period 22:40–00:30 of 13 August.

near the ground surface to values less than 1 mHz at the heights between 1- and 3 km. Despite a presence of vertical oscillations in the profile of $N(z)$, the value of $N(z)$ averaged over the thickness of the lower 1-km atmospheric layer is higher than the average value of N in the layer between 1 and 3 km. Such a stratification of $N(z)$ leads to the formation of a wave guide for internal waves in the lower atmospheric layer with a height of about 1 km.

The obtained theoretical form for the frequency spectrum $S_{1E}(\omega)$ of the vertical displacements of fluid parcels in the Eulerian frame may be presented as a sum of two terms (Chunchuzov, 2003, Chunchuzov *et al.*, 2004): $S_{1E}(\omega) = S'_{1E}(\omega) + S''_{1E}(\omega)$, where $S'_{1E}(\omega)$ and $S''_{1E}(\omega)$ are the contributions to the frequency spectrum from the low-wavenumber and high-wavenumber parts of the 4D frequency-wavenumber spectrum, respectively. Similarly for the frequency spectrum of the horizontal wind speed fluctuations, we have $V_{1E}(\omega) = V'_{1E}(\omega) + V''_{1E}(\omega)$.

In the range $\Omega_0^2 \ll \omega^2 < N_1^2$ the obtained approximate forms for the terms $S'_{1E}(\omega)$ and $V'_{1E}(\omega)$ are given by expressions (Chunchuzov *et al.*, 2004)

$$S'_{1E}(\omega) \approx \nu_0^2 \Omega_0 / (2\omega^2), \quad V'_{1E}(\omega) \approx \epsilon / (2\omega^2),$$

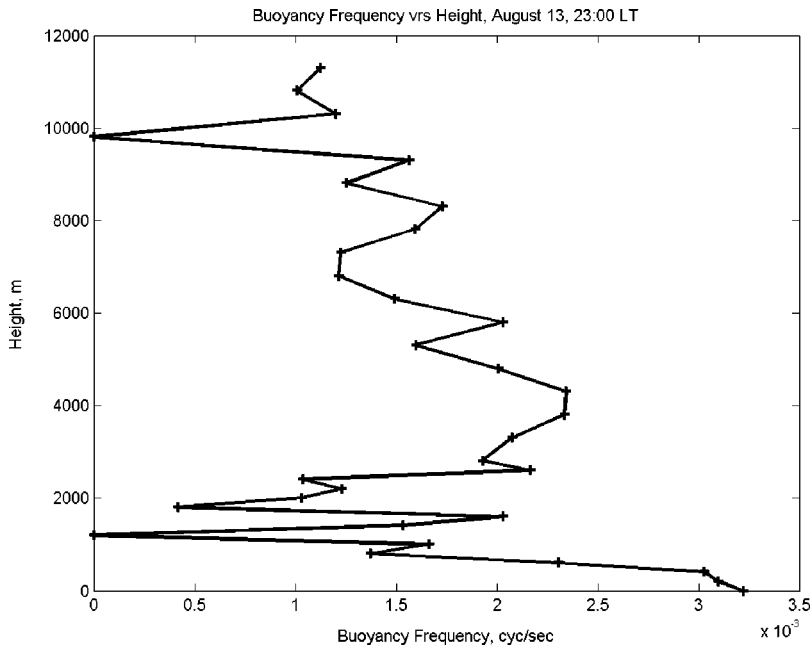


FIG. 11. BV-frequency profile calculated from the radiosonde measurements of the temperature in the troposphere, 13 August 2002, 23:00.

$$\epsilon = \sigma^2 \Omega_0, \quad (4)$$

where ν_v is the rms value of the vertical parcel displacements caused by a random internal wave field, $\sigma^2 = N_1^2 \nu_v^2$ is the mean square value of the wind speed fluctuations, and Ω_0 is the characteristic frequency of the internal wave field. This frequency is connected by a dispersion equation $\Omega_0 = N_1 k_0 / m_0$ to the horizontal and vertical wavenumbers, k_0 and m_0 , of the maximum of the 3D wave number spectrum of the vertical displacements. Hence Ω_0 may be treated as the frequency for which a maximum of wave energy is “pumped” from random internal wave sources to the wave system.

The ratio between the characteristic wave numbers $\chi \equiv k_0 / m_0$ is of the same order as the ratio σ_\perp / σ_v , where σ_\perp and σ_v are the rms values of the fluctuations of horizontal and vertical wind velocity, respectively. Under stable stratification, vertical velocities are significantly suppressed, therefore $\chi \gg 1$, and $\Omega_0^2 \ll N_1^2$. According to (4) the coefficient ϵ is the product of the rms velocity σ of the parcels and their acceleration $\sigma \Omega_0$, where $\Omega_0 = N_1 / \chi$. Therefore ϵ may be treated as the mean rate of wave energy input from the random internal wave sources.

For comparison with the experimental data we have to use the one-sided frequency spectra defined as $2S_{1E}(\omega)$ and $2V_{1E}(\omega)$ for $\omega \geq 0$ and 0 for $\omega < 0$. These spectra decay as ω^{-2} in the range $\Omega_0^2 \ll \omega^2 < N_1^2$. However, as ω approaches the values close to N_1 the spectrum $2S_{1E}(\omega)$ shows a local maximum [Fig. 12(a)]. This maximum is caused by a constructive interference between the incident and reflected internal waves from the jump of the BV-frequency value at $z = h$. When ω tends to N_1 , the parcel displacements induced by trapped IGWs become almost vertical, whereas horizontal displacements and velocities tend to zero (Gossard and Hooke, 1975).

At high vertical wave numbers, the horizontal phase speeds of the internal wave modes are comparable to the values of σ , therefore the nonlinear interactions (both reso-

nant and Γ_A nonresonant) between these modes play a key role in the formation of the IGW spectrum. These interactions generate in the frequency-wavenumber spectrum of the internal wave field the spectral components with high vertical wave numbers and high frequencies (Churchuzov, 2002, 2003). Particularly, the nonlinear interactions between any two spectral components with the frequencies $\omega_1 < N_1$ and $\omega_2 < N_1$ can generate the new spectral components with the combinative frequencies $\omega_1 + \omega_2 > N_1$. These interactions do not necessarily satisfy resonant conditions $\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}$, $\omega_1 + \omega_2 = \omega$, where $\omega = \omega(\mathbf{k})$ is the dispersion relation for the linear internal wave modes, and $\omega_1 = \omega(\mathbf{k}_1)$, $\omega_2 = \omega(\mathbf{k}_2)$. Therefore, the nonresonant interactions are not selective ones: each wave mode may interact with any other mode, generating new modes so that a wave energy cascade is developing. This cascade leads to the wave energy transfer from the frequency region $\omega < N_1$ to the high-frequency region $\omega > N_1$, thereby generating a high-frequency tail in the frequency internal wave spectrum. This tail is similar in its origin to the k^{-3} -tail in the vertical-wave number spectrum of internal wave field (Hines, 2001; Churchuzov, 2002).

The high-frequency parts of the one-sided frequency spectra $2S_{1E}''(\omega)$ and $2V_{1E}''(\omega)$, generated by the nonlinearity of the internal wave field, is described by the following expressions (Churchuzov *et al.*, 2004):

$$2S_{1E}''(\omega) \approx \beta \sigma_v^2 / \omega^3, \quad 2V_{1E}''(\omega) \approx \alpha N_1^2 \sigma_v^2 / \omega^3 \quad (\omega^2 \gg \Omega_0^2). \quad (5)$$

Here, α and β are the numerical coefficients depending on the mean square vertical gradient $\langle (\partial s_3 / \partial z)^2 \rangle$ of the vertical displacements s_3 caused by IGWs, and σ_v^2 is the mean square value of the vertical wind velocity fluctuations. The magnitudes of α and β increase with a growth of the value of $\langle (\partial s_3 / \partial z)^2 \rangle$ up to some maximum of the order of 0.1, which characterizes a nonlinear saturation of the amplitude of the spectral tail. Thus a strong nonlinearity of the internal wave field forms a universal high-frequency tail with a -3

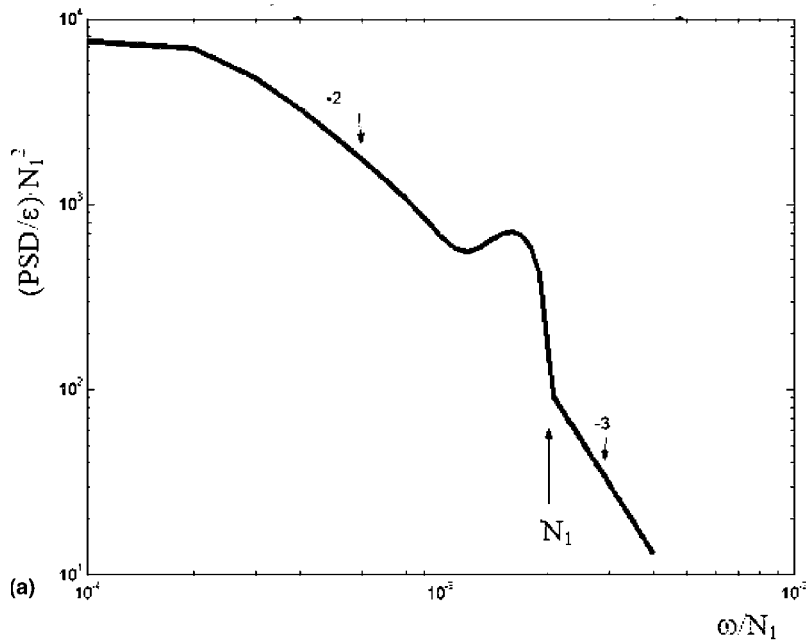
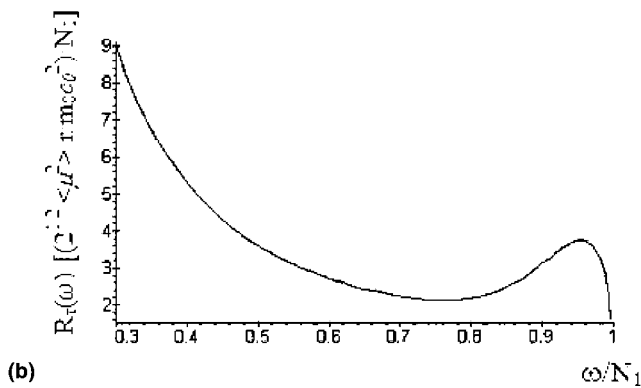


FIG. 12. Frequency spectra of the vertical displacements and travel time fluctuations calculated from the model of internal wave spectrum in stably stratified atmosphere. (a) Frequency spectrum of the vertical displacements defined by expressions (4) and (5) and calculated for $m_0 z = 10$, $N_1 = 0.001$ Hz, $\chi = 11.5$, $\epsilon = 10^{-5}$ $\text{m}^2 \text{s}^{-3}$, $\beta = 0.23$; $\sigma_v = 0.03$ m/s. (b) Normalized frequency spectrum of travel time fluctuations.



power law decay. This tail, shown in Fig. 12(a), becomes important at frequencies $\omega > N_1$.

B. Model of the frequency spectrum of travel time fluctuations for the acoustic pulses propagating through a random internal wave field

The knowledge of the frequency-wave number spectrum $G_\mu[\mathbf{k}, \omega, z]$ of the acoustic refractive index $\mu(\mathbf{x}, t)$ in a moving inhomogeneous medium allows one to calculate the form of the frequency spectrum $R_T(\omega)$ of the pulse travel time fluctuations $\delta T(t)$ along a selected ray trajectory Γ (Chunchuzov, 2003):

$$R_T(\omega) = 2\pi c_0^{-2} \int_{\Gamma} dx \int dk_3 \int dk_2 G_\mu \times [-k_3 \tan \theta(x), k_2, k_3, \omega, z(x)], \quad (6)$$

where $\mathbf{k} = (k_1, k_2, k_3)$ is the wave number vector, x is the horizontal distance from a point acoustic source, $z = z(x)$ is the function that describes a form of the refractive ray trajectory Γ in the acoustic wave guide caused by the mean temperature and wind speed stratification near the ground

surface, and $\theta(x)$ is the elevation angle of the ray trajectory at some point $(x, z(x))$. The calculation of the statistics of travel time fluctuations is similar to those in the ocean (Flatté *et al.*, 1979), but here we take into account the wind speed fluctuations, which play a substantial role in the atmosphere.

As shown by Chunchuzov (2003) the spectrum $G_\mu[\mathbf{k}, \omega, z]$ for a random internal wave field may be expressed as a sum,

$$G_\mu[\mathbf{k}, \omega, z] = S_{4E}(\mathbf{k}, \omega, z) (N^2 g^{-1})^2 / 4 + c^{-2} G_{11}(\mathbf{k}, \omega, z), \quad (7)$$

in which the first term describes a spectral contribution coming from the relative temperature fluctuations T'/T_0 (they are proportional to the vertical parcel displacements), g is the gravity acceleration, and second term $G_{11}(\mathbf{k}, \omega, z)$ is the 4D spectrum of the fluctuations of the horizontal wind speed projection v_x along a ray trajectory Γ .

The approximate forms for the frequency wave number spectra $S_{4E}(\mathbf{k}, \omega, z)$ and $G_{11}(\mathbf{k}, \omega, z)$ have been provided in Chunchuzov (2002). By substituting these forms into the expression (7), and taking this integral in the case of small elevation angles [$\theta(x) \approx 0$], one obtains the form of the

FREQUENCY SPECTRA OF THE EFFECTIVE SOUND SPEED FLUCTUATIONS
IN STABLE ABL, Aug 10-13, 2002

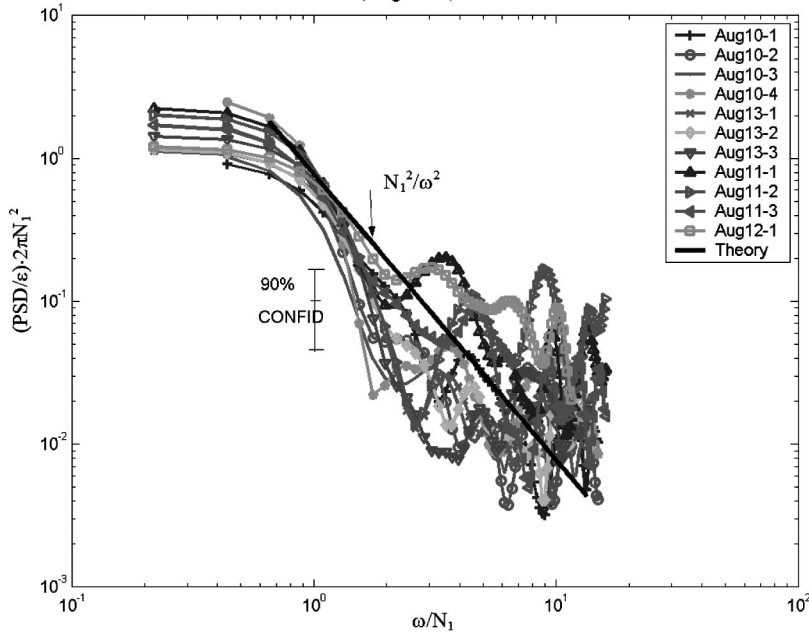


FIG. 13. Normalized frequency spectra of the pulse travel time (or effective sound speed) fluctuations measured at the points R1–R4 during 10–13 August 2002. Horizontal axis: frequency normalized by $N_1=0.001$ Hz; vertical axis: frequency spectrum normalized by ϵ/N_1^2 . Theoretical spectrum $\epsilon\omega^{-2}$ is shown by a straight continuous line.

travel time spectrum $R_\tau(\omega)$ shown in Fig. 12(b) [taken from Churchuzov (2003)]. In this figure the frequency ω is normalized by the BV frequency N_1 (horizontal axis), whereas the spectrum $R_\tau(\omega)$ is normalized by $[(\sqrt{2}\langle\mu^2\rangle r/m_0c^2)/N_1]$, which is proportional to the mean square value of the pulse travel time fluctuations at a distance $x=r$, where $\langle\mu^2\rangle\approx\langle T'^2/4T_0^2\rangle+\langle v_x^2/c^2\rangle$ is the mean square value of the acoustic refractive index fluctuations in the moving atmosphere. This spectrum decays as ω^{-2} with an increase of ω , but reaches a local maximum near $\omega\approx N_1$ similarly to the frequency spectrum of the vertical displacements.

C. Comparison between theoretical and observed frequency spectra

The frequency spectra for the effective sound speed fluctuations $\langle(\delta c_{\text{eff}})_m(t)\rangle$ near ray turning points obtained during the period from 10 to 13 August 2002 are shown in Fig. 13. These spectra are calculated from the time series of the acoustic pulse travel time fluctuations measured at the points R1–R4 (shown in Fig. 1) and are normalized by the value of ϵ/N_1^2 , where N_1 is the mean value of the BV frequency in the lower atmospheric layer with a thickness of 1 km. According to the radiosonde data during measurements of the fluctuations $(\delta c_{\text{eff}})_m(t)$, this value ranged from 1 to 2 mHz.

The value of $\epsilon=\sigma^2 N_1/\chi$ was estimated from the rms value of the measured fluctuations $\langle(\delta c_{\text{eff}})_m(t)\rangle$, which are of the same order as the rms value of the wind speed fluctuations σ . The ratio $\chi=\sigma_\perp/\sigma_v$ was estimated from the obtained time series of the fluctuations of horizontal wind speed components, v_x , v_y , and vertical component, w , by calculating the ratio between their rms values σ_\perp and σ_v . For instance, during 10 August 2002, $\sigma\approx 0.12$ at the height of 99 m, and $\chi=\sigma_\perp/\sigma_v\approx 11.5$, so ϵ is about $10^{-5}\text{ m}^2\text{ s}^{-3}$.

As shown in Fig. 13 the obtained frequency spectra decrease with increasing normalized frequency f/N_1 in the range 0.6–2, where $N_1=1\times 10^{-3}$ Hz (the corresponding periods range from 28 to 8 min). At higher relative frequencies f/N_1 , between 2 and 10 (the corresponding periods are between 8 and 1.5 min), the spectral density has a number of local peaks. Note that for the frequencies more than 2 mHz, internal waves are trapped in the lower layer with the height of a few hundred meters, where a static stability $N(z)$ reaches a maximum [Fig. 9(b)]. As seen from Fig. 13, the frequencies of the spectral peaks and the distribution of the spectral density over frequencies vary slightly from case to case, but despite such variability of the observed spectra, their mean slopes and amplitudes are close to those predicted theoretically.

IV. CONCLUSIONS

The method of acoustic pulse sounding of stable ABLs allowed us to obtain new data about variances, coherences, and horizontal speeds of propagation of meso-scale wind speed fluctuations in the stable ABL. Several dominant periods were found in the coherences and frequency spectra of the wind speed and pulse travel time fluctuations. These periods, and the corresponding horizontal speeds and scales of the observed fluctuations, are typical for the internal gravity waves in the lower atmosphere. Among these waves are both the trapped modes in the lower part of the ABL with a maximum static stability and the freely propagating modes, which may be generated by the tropospheric internal wave sources such as unstable wind shears. The existence of the wave guide near the ground surface was confirmed by the analysis of the vertical profiles of the horizontal wind speed and temperature (or BV frequency) in the stable ABL. The latter were continuously monitored by a Doppler sodar and temperature profiler.

Based on the model of an internal wave spectrum in the stably stratified atmospheric layer we obtained theoretical forms for the frequency spectra of both the wind speed and pulse travel time fluctuations induced by internal waves. Despite a temporal variability of the obtained experimental spectra, their average power law decay was close to the predicted theoretical power law $\epsilon\omega^{-2}$, where ϵ is the mean rate of wave energy input from the random internal wave sources. The value of ϵ obtained from the experimental data gives a satisfactory estimate of the spectral density of meso-scale wind speed fluctuations in a stable ABL.

ACKNOWLEDGMENTS

This work was supported by Russian Foundation of Basic Research, Grant Nos. 03 05-04001 and 02-05 65112.

- Aggarwal, S. K., Singal, S. P., and Sristava, S. K. (1980). "Sodar studies of gravity waves in the planetary boundary layer at Delhi," *Mausam* **31**, 373–384.
- Allen, K. A., and Joseph, R. I. (1989). "A canonical statistical theory of oceanic internal waves," *J. Fluid Mech.* **204**, 185–228.
- Anderson, P. S. (2002). "Fine scale structure observed in stable boundary layers by sodar and tethered sonde," *Proc. of the 11th Intern. Symp. on Acoustic Remote Sensing (ISARS'2002)*, Rome, Italy, pp. 409–412.
- Avilov, K. V. (1995). "Pseudo-differential parabolic equations of sound propagation in the ocean with smooth horizontal inhomogeneities, and their numerical solutions," *Akust. Zh.* **41**(1), 5–12 (in Russian).
- Bendat, J. S., and Piersol, A. G. (1967). *Measurement and Analysis of Random Data* (Wiley, New York).
- Blumen, W., Banta, R., Durms, S. P., Fritts, D. C., Newsom, R., Poulos, G. S., and Sun, J. (2001). "Turbulence statistics of a Kelvin-Helmholtz billow event observed in the nighttime boundary layer during the CASES-99 field program," *Dyn. Atmos. Oceans* **34**, 189–204.
- Chernov, L. A. (1960). *Wave Propagation in a Random Medium* (Dover, New York).
- Chimonas, G. (1978). "Packet-scale motions forced by nonlinearities in a wave system," *J. Atmos. Sci.* **35**, 382–395.
- Chimonas, G. (1999). "Steps, Waves and Turbulence in the Stably Stratified Planetary Boundary Layer," *Boundary-Layer Meteorol.* **90**, 398–416.
- Chunchuzov, I. P. (1996). "The spectrum of high-frequency internal waves in the atmospheric waveguide," *J. Atmos. Sci.* **53**, 1798–1814.
- Chunchuzov, I. P. (2002). "On the high-wavenumber form of the Eulerian internal wave spectrum in the atmosphere," *J. Atmos. Sci.* **59**, 1753–1772.
- Chunchuzov, I. P. (2003). "Influence of internal gravity waves on sound propagation in the lower atmosphere," *Meteorol. Atmos. Phys.* **34**, 1–16.
- Chunchuzov, I. P., Bush, G. A., and Kulichkov, S. N. (1990). "On acoustical impulse propagation in a moving inhomogeneous atmospheric layer," *J. Acoust. Soc. Am.* **88**, 455–466.
- Chunchuzov, I. P., Otrezov, A. I., and Petenko, I. V. (1996). "On the inclined sounding of the stably stratified atmospheric boundary layer," *Proc. of the 8th Int. Symp. on Acoustic Remote Sensing (ISARS'96)*, Moscow, pp. 65–71.
- Chunchuzov, I. P., Otrezov, A. I., Petenko, I. V., Tovchigrechko, V. N., Svertilov, A. I., Fogel, A. L., and Fridman, V. E. (1997). "Travel time fluctuations of the acoustic pulses propagating in the atmospheric boundary layer," *Izv., Acad. Sci., USSR, Atmos. Oceanic Phys.* **3**, 324–338.
- Chunchuzov, I. P., Kulichlov, S. N., Otrezov, A. I., Perepelkin, V., Kallistratova, M. A., Tovchigrechko, V. N., Kadygrov, E. N., and Kuznetsov, R. (2004). "Study of the spectral characteristics of the meso-scale fluctuations of wind speed and acoustic signal travel time in stably stratified atmospheric boundary layer," *Izv. Acad. Sci., USSR, Atmos. Oceanic Phys.* (in press).
- Danilov, S. D., and Chunchuzov, I. P. (1992). "Possible Mechanism of Layered Structure Formation in a Stably Stratified Atmospheric Boundary Layer," *Izv., Acad. Sci., USSR, Atmos. Oceanic Phys.* **28**, 684–688.
- Derbyshire, S. H. (1999). "Stable boundary-layer modelling: established approaches and beyond," *Boundary-Layer Meteorol.* **90**, 423–446.
- Einaudi, F., and Finnigan, J. J. (1993). "Wave-turbulence dynamics in the stably stratified boundary layer," *J. Atmos. Sci.* **50**, 1841–1863.
- Flatté, S., Dashen, R., Munk, W., Watson, K., and Zachariassen, F. (1979). *Sound Transmission through a Fluctuating Ocean* (Cambridge U.P., Cambridge).
- Galaktionov, A., and Maderich, V. (1999). "Dynamics of the layered structures in the final stage of turbulence decay in the stably stratified fluid," *Izv., Acad. Sci., USSR, Atmos. Oceanic Phys.* **35**, 829–837.
- Gossard, E. E., and Hooke, W. H. (1975). *Waves in the Atmosphere* (Elsevier, Amsterdam).
- Gossard, E. E., Gaynor, J. F., Zamor, R. J., and Neff, W. D. (1985). "Fine structure of elevated stable layers observed by sounder and in situ Tower Sensors," *J. Atmos. Sci.* **42**, 2156–2169.
- Grachev, A. I., Zagoruiko, S. V., Matveev, A. K., and Mordukhovich, M. I. (1978). "Some results of recording of atmospheric infrasound waves," *Izv., Acad. Sci., USSR, Atmos. Oceanic Phys.* **14**, 474–482.
- Hines, C. O. (2001). "Theory of the Eulerian tail in the spectra of atmospheric and oceanic internal gravity waves," *J. Fluid Mech.* **448**, 289–313.
- Kjelaas, A. G., Beran, D. W., Hooke, W. H., and Bean, B. R. (1974). "Waves Observed in the Planetary Boundary Layer using an Array of Acoustic Sounders," *J. Atmos. Sci.* **31**(8), 2040–2045.
- Lighthill, J. (1978). *Waves in Fluids* (Cambridge U.P., Cambridge).
- Ostashev, V. E., Brahler, B., Mellert, V., and Goedecke, G. H. (1998). "Coherence functions of plane and spherical waves in a turbulent medium with the von Karman spectrum of medium inhomogeneities," *J. Acoust. Soc. Am.* **104**, 727–737.
- Tatarskii, V. I. (1971). "The Effects of the Turbulent Atmosphere on Wave Propagation," (US Dept. Commerce, Springfield).
- Weinstock, J. (1984). "Effect of gravity waves on turbulence decay in stratified fluids," *J. Fluid Mech.* **140**, 11–26.
- Wilson, D. K., and Thomson, D. W. (1994). "Acoustic propagation through anisotropic, surface-layer turbulence," *J. Acoust. Soc. Am.* **96**, 1080–1095.
- Wilson, D. K., Ziemann, A., and Ostashev, V. E. (2001). "An overview of acoustic travel-time tomography in the atmosphere and its potential applications," *Proc. of the Intern. Workshop "Tomography and Acoustics: Recent developments and methods"*, March 2001, Institute for Meteorology, Leipzig, Germany, pp. 94–101.

The use of impedance matching capillaries for reducing resonance in rosette infrasonic spatial filters

Michael A. H. Hedlin^{a)}

*Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography,
University of California, San Diego, La Jolla, California 92093-0225*

Benoit Alcoverro^{b)}

*Commissariat à l'Énergie Atomique, Département Analyse et Surveillance de l'Environnement, BP 12,
91680 Bruyères le Chatel, France*

(Received 19 May 2003; revised 2 March 2004; accepted 1 April 2004)

Rosette spatial filters are used at International Monitoring System infrasound array sites to reduce noise due to atmospheric turbulence. A rosette filter consists of several clusters, or rosettes, of low-impedance inlets. Acoustic energy entering each rosette of inlets is summed, acoustically, at a secondary summing manifold. Acoustic energy from the secondary manifolds are summed acoustically at a primary summing manifold before entering the microbarometer. Although rosette filters have been found to be effective at reducing infrasonic noise across a broad frequency band, resonance inside the filters reduces the effectiveness of the filters at high frequencies. This paper presents theoretical and observational evidence that the resonance inside these filters that is seen below 10 Hz is due to reflections occurring at impedance discontinuities at the primary and secondary summing manifolds. Resonance involving reflections at the inlets amplifies noise levels at frequencies above 10 Hz. This paper further reports results from theoretical and observational tests of impedance matching capillaries for removing the resonance problem. Almost total removal of resonant energy below 5 Hz was found by placing impedance matching capillaries adjacent to the secondary summing manifolds in the pipes leading to the primary summing manifold and the microbarometer. Theory and recorded data indicate that capillaries with resistance equal to the characteristic impedance of the pipe connecting the secondary and primary summing manifolds suppresses resonance but does not degrade the reception of acoustic signals. Capillaries at the inlets can be used to remove resonant energy at higher frequencies but are found to be less effective due to the high frequency of this energy outside the frequency band of interest. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1760778]

PACS numbers: 43.28.Dm, 43.28.Ra, 43.50.Cb [LCS]

Pages: 1880–1888

I. INTRODUCTION

A. Infrasonic and global nuclear monitoring

The Comprehensive Nuclear Test-Ban Treaty (CTBT) was made available for signature at the United Nations in September of 1996. This has led to increased interest in monitoring globally for nuclear testing activity at all yields which in turn has led to the development of the four-network International Monitoring System (IMS). The IMS includes infrasound, radionuclide, seismic, and hydroacoustic stations. Although the primary goal of the system is to monitor Earth's atmosphere, solid interior, and oceans for nuclear testing activity, an additional benefit of the continuously monitoring system is that it will allow detection, tracking, and analysis of other man-made and natural phenomena on an unprecedented global scale.

This paper is concerned with the infrasound component of the IMS (Christie *et al.*, 2001). As of November 2001, 10 of the planned 60 microbarograph arrays were operating and another 10 were under construction. Completion of the entire network is expected to take several more years. Each infra-

sound array will comprise four to eight sensors with a baseline of 1 to 3 km. Sensors in the infrasound network are designed to detect signals between 0.01 and 8 Hz, but the array geometry is designed for optimal detection in the 0.1–1 Hz range. Infrasonic pressure data are collected at each site at 20 samples per second (sps) along with wind speed and direction sampled at 1 sps. The operating arrays transmit data to the International Data Center (IDC) in Vienna, Austria. One array (Fig. 1) is in the Anza-Borrego desert south of Palm Springs, CA.

B. Reduction of infrasonic noise

The principal source of noise in the frequency band of interest to the nuclear monitoring community is atmospheric turbulence. As a result, all pressure data are filtered in the field by a spatial wind-noise-reducing system (Daniels, 1959; Burrige, 1971; Grover, 1971). The filter averages out pressure variations that are due to local turbulence and increases the signal-to-noise ratio (SNR).

The infrasound monitoring network is designed to provide monitoring of the atmosphere at all points on the globe. As a result, the noise reducing systems are required to be omnidirectional. A recent design of a spatial filter by Alcov-

^{a)}Electronic mail: hedlin@ucsd.edu

^{b)}Electronic mail: benoit.alcoverro@cea.fr

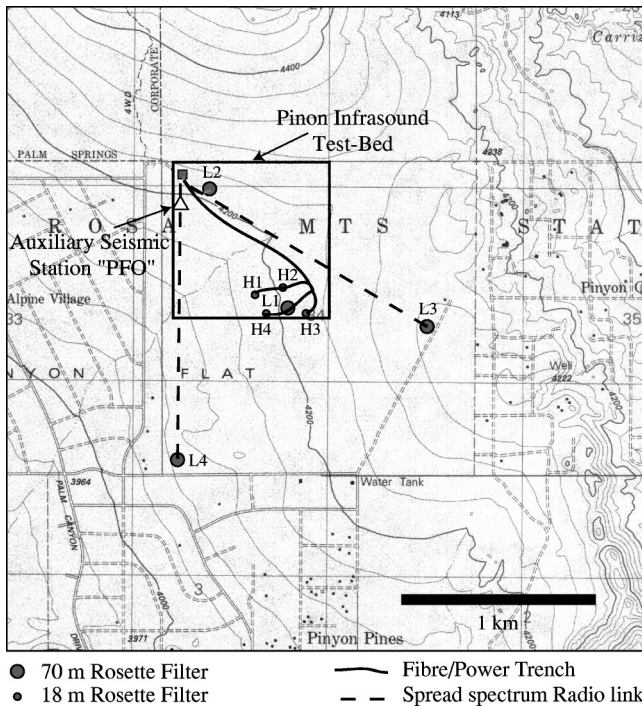


FIG. 1. The eight-element infrasound array at Pinon Flat, CA comprises four long-period elements (L1–L4) in a centered triangle spanning less than 2 km and four short-period elements (H1–H4) in an irregular quadrilateral near the center. The terrain slopes down gently to the southwest. The contour interval is 40 ft. (12.2 m).

erro (1998) consists of several clusters, or rosettes, of low-impedance inlets. Low-impedance inlets are preferred as they require essentially no maintenance. Each inlet in a cluster is connected to a “secondary” summing manifold by a solid pipe. A solid pipe connects each secondary summing manifold to one primary summing manifold or directly to the sensor. All pipes are solid, and thus all acoustic energy that propagates to the sensor enters the filter via the inlets. The signals are integrated acoustically at the summing manifolds. Christie (Provisional Technical Secretariat, Comprehensive Nuclear-Test-Ban Treaty Organization) has proposed several modifications to Alcoverro’s original rosette design (Christie, 1999). The Christie rosette filters range from 92 ports distributed over an area 18 m across to 144 ports spanning an area 70 m across (Fig. 2). The smaller filter is considered to be more effective at high frequencies at windy sites. The 70-m aperture filter is considered appropriate for windy sites that require suppression of noise between 0.02 and 0.7 Hz.

C. Resonance in rosette filters

As discussed by Hedlin *et al.* (2003) and Alcoverro and LePichon (2004), one drawback of the rosette filter design currently in use at IMS array sites is internal resonance. Impedance discontinuities exist at the inlets to the summing manifolds, the microbarometer, and the inlets. The resonance is predicted and observed above 3 Hz in data from the 18-m rosette systems (Fig. 3). Resonance peaks exist above 0.7 Hz and peaks at 2.65 and 7.95 Hz in data from the 70-m rosette systems (Fig. 3). Due to attenuation, the resonant energy is seen in data from the 70-m rosettes at all frequencies above 0.7 Hz. Higher overtones are believed to exist at higher fre-

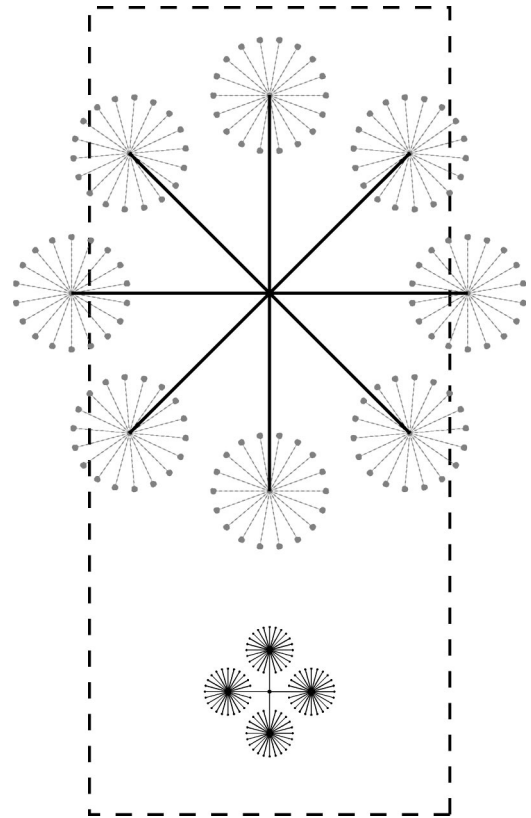


FIG. 2. Two rosette filters considered in this paper are shown to scale with a National Football League playing surface (dashed rectangle). The playing surface is 50 m wide by 112 m long. The 18-m filter comprises 92 low-impedance inlets in four rosettes. The 70-m filter comprises 144 inlets arranged in eight rosettes. In both filter designs, the inlets are connected by solid pipe to secondary summing manifolds at the center of each rosette. The secondary manifolds are connected to a primary manifold, which is connected by a short pipe to the sensor. Each inlet is located the same distance along solid pipe from the sensor. Signals arriving from directly above the filter are summed in phase. Adjacent ports in the 70-m filter are separated by 2.79 m. Adjacent ports in the 18-m filter are 0.85 m apart. For this study, we used the filters at locations L2 and H1 in the array IS57. The first experiment involved simplified rosette filters consisting of eight 27 m pipes attached to a primary summing manifold. These filters are shown to scale in this figure as the solid black lines. The inlets in some of the filters were fitted with capillaries.

quencies but are removed in our data by the anti-aliasing filter, which takes effect at 8 Hz. The resonance is observed at all wind speeds (Hedlin *et al.*, 2003; Alcoverro and LePichon, 2004). The resonance results in an increase in noise levels and the signal waveform may be significantly distorted.

D. Rationale for this study

There are two primary objectives for this study. The first is to identify the reflection points that give rise to resonant energy in the passband below 8 Hz in the 18- and 70-m rosette filters. Toward this end we present data collected by simplified rosette filters that comprise just eight inlets and span 54 m (Fig. 2). We use the theory of Alcoverro and LePichon (2004) to model the frequency response of the simplified and full rosette filters. The second, and key, objective of this paper is to investigate the use of impedance matching capillaries to suppress, or remove, resonance from these fil-

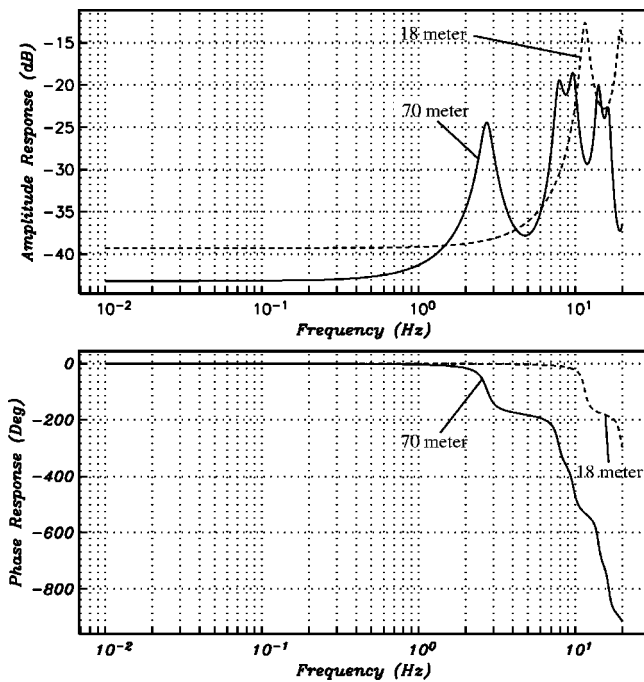


FIG. 3. Predicted amplitude and phase response of the 18-m/92-rosette filter (dashed curve) and the 70-m/144-rosette filter (solid curve) for one inlet. The resonance peaks coincide with significant change in the phase response of the filter. The long-period response is given by $-20 \log_{10}(N)$, where N is the number of inlets.

ters. We predict theoretically the response of simple and full rosette filters equipped with impedance matching capillaries and compare the theoretical results with recorded data. We use continuous recordings of atmospheric noise and a large acoustic signal to validate the theoretical results.

II. THE PINON FLAT INFRASOUND TEST-BED

All of our experiments have been conducted at the infrasound test-bed at the Pinon Flat Observatory in southern California. The observatory is described in detail by Hedlin *et al.* (2003) and so only a brief description is given here. The observatory is located in the desert south west of Palm Springs. This site has been useful for studies of infrasonic wind-noise reducing systems as wind speeds vary from near zero at night to above 15 m/s during the day. The observatory is sparsely vegetated with Pinon pine trees and there is essentially no ground cover. The IMS infrasound array IS57 is located at the observatory (Fig. 1). The array comprises eight sensors on a baseline of 2 km and uses both the high-frequency, 18-m, and low-frequency, 70-m, rosette filters described earlier.

III. EXPERIMENTAL LAYOUT

A typical experimental layout is as described by Hedlin *et al.* (2003). Each site includes an MB2000 aneroid microbarometer. The pressure data are low-pass filtered below 8 Hz and sampled at 20 sps by a Reftek digitizing system. Air temperature and wind velocity were collected at 1 sps at heights of 1 and 2 m, respectively. Data from the temporary sites were transmitted in real-time to our laboratory in La Jolla via a 2.2-GHz spread-spectrum radio link. All recording

devices were housed in insulating cases to shield them from the extreme desert conditions. In all experiments, one microbarograph sensor was attached to a single, low-impedance, inlet located 5 cm above the ground and was used to indicate the level of background noise. All other sensors were attached to a noise-reducing system. With the exception of two sites, all systems were deployed temporarily. Two of the sites used for this study are in the IMS infrasound array IS57. Data from the array sites were transmitted to our laboratory via phone lines.

IV. PRELIMINARY TESTS OF CAPILLARIES

A. Observed and predicted resonance

The infrasonic noise-reducing filter developed by Daniels in the late 1940s comprised a single pipe with inlets distributed along its length (Daniels, 1959). The prototype was 603 m long with 100 inlets. The pipe diameter, and thus its characteristic impedance, changed, in steps, at each of the inlets, the section increasing toward the sensor. The acoustic impedance of each inlet was selected in the aim that the parallel combination of the inlet and the characteristic impedance of the small pipe was equal to the characteristic impedance of the larger pipe at the location of the inlet. Acoustic energy propagating in the pipe away from the sensor would not reflect back to the sensor, but impedance and time lag at each inlet are different.

As shown in Fig. 2, a rosette filter consists of inlets connected via solid pipes and summing manifolds to a microbarograph located at the center. In the 70-m rosette filters in use at the IMS array IS57, the pipes connecting the central, primary, summing manifold with the secondary summing manifolds are 27 m long, with an internal diameter of 2.1 cm. The characteristic acoustic impedance of these pipes is 1.202 mega-ohms ($M\Omega$), where acoustic resistance is calculated using SI units (i.e., Ns/m^5). The secondary summing manifolds are connected via 8-m-long pipes to the inlets. These pipes have an inner diameter of 1.6 cm and a characteristic impedance of 2.07 $M\Omega$. As discussed by Hedlin *et al.* (2003), resonance peaks are observed at 2.65 and 7.95 Hz. Resonance modeling reproduces these resonance peaks (Fig. 3) and indicates that the reflections that give rise to these peaks occur at the secondary and primary summing manifolds. Data from the 18-m rosette filters do not exhibit a prominent spectral peak due to resonance, but a gradual increase in energy with increasing frequency above 2 Hz. Peaks above the Nyquist frequency at 10 Hz are predicted (Fig. 3).

To verify that the reflections that produce significant resonance peaks below 10 Hz are caused by reflections at the summing manifolds, and not at the inlets, and that capillaries can remove the resonant energy without also attenuating the signals of interest, we conducted a field experiment with simple, eight-arm, filters. As shown in Fig. 2, we constructed filters that spanned 54 m, or exactly the area inside the secondary summing manifolds in a 70-m rosette filter. A simple, eight-arm, filter is depicted in Fig. 2 as the solid black lines.

The experiment with simple filters occurred from 1 November, 2001 to 14 January, 2002. For this experiment, we

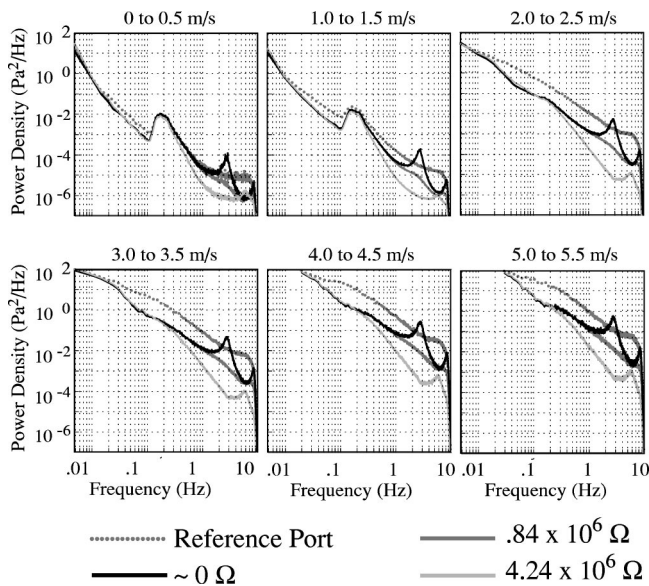


FIG. 4. Robust estimates of the dependence of infrasonic noise on wind speed and frequency can be obtained by stacking spectral estimates taken at different times but during similar wind conditions. To accomplish this, we associate each spectral estimate with a single wind-speed value: the average wind speed from the 15-m interval spanned by the spectrum. We stack spectral estimates after binning them by wind speed into bins that span 0.5 m/s starting at 0.0 m/s. The six panels in this figure show binned and stacked spectra taken at the reference port (dotted) and three eight-arm simple “rosette” filters. The filters were identical except for the acoustic resistances installed at the inlets. The black curves represent data from the filter with no acoustic resistance installed at the inlets. The dark gray and light gray curves represent data from filters with intermediate and high acoustic impedance capillaries installed at the inlets.

co-located three simple filters with a reference site. The simple filters used in this test were identical, except at the inlets. One filter (black curves in Fig. 4) was equipped with low-impedance inlets. The other two filters were equipped with capillaries at each of the inlets. One filter (dark gray curves) included acoustic resistances at the inlets of 0.84 MΩ. The other (light gray curves) was equipped with higher acoustic resistance at the inlets of 4.24 MΩ. These resistances were selected to bracket the characteristic impedance of the 27-m pipes connecting the inlets with the primary summing manifold and the microbarograph at each site.

Spectra from the four sites are displayed in Fig. 4. In this figure, we show binned and stacked spectral density from the four sites at wind speeds ranging from 0–0.5 m/s to 5.0–5.5 m/s. In each panel, the dotted curves represent ambient noise levels recorded at the reference site. The growth of infrasonic noise levels with increasing wind speed is evident. As shown in Fig. 4, the system with open inlets produces a substantial resonance peak in the data at 2.65 Hz at all wind speeds. The first resonance peak in data from the full 70-m rosette filters was also observed at 2.65 Hz (Hedlin *et al.*, 2003). This provides empirical evidence of the finding made by Hedlin *et al.* (2003) that the reflections that give rise to the resonance at this frequency are occurring at the secondary summing manifolds, and not at the inlets. The lowest resonant frequency is determined by the propagation time in the longer (27 m) pipe in the noise filter.

As shown in Fig. 4, it is possible to blunt, or entirely remove, the resonance peak at 2.65 Hz by adding acoustic

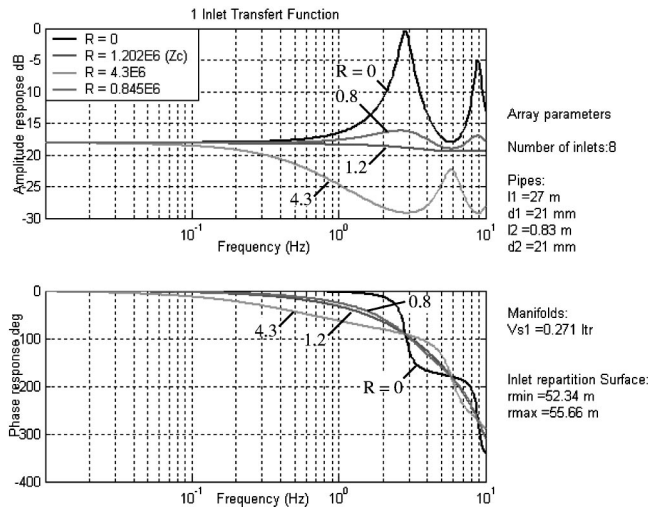


FIG. 5. Theoretical predictions of amplitude and phase response for the eight-arm “rosette” filter systems. The theoretical predictions match closely the spectra obtained from recorded data.

resistance at the inlets. The system with inlet resistance that is less than the characteristic impedance of the pipe shows some evidence of resonance. The system with the high inlet resistance shows a marked decrease in noise levels at 0.3 Hz with no resonance peak at 2.65 Hz. A resonance peak is observed at 5.3 Hz, or exactly double the frequency of the unmodified system. The spectral results indicate that resonance can be controlled by adding acoustic resistance at the inlets, but suggests that high resistance will attenuate the energy across a broad band.

The response of the eight-arm “simple” filters has been calculated using the method of Alcoverro and Le Pichon (2004). Four simple filters were considered. The first three were exactly as tested in the field. The fourth filter included acoustic resistance at the inlets of 1.202 MΩ, equal to the characteristic impedance of the 27-m pipes leading to the primary summing manifold. The capillaries are modeled as a pure resistance with a value depending on the diameter and the length of the small pipe used. These elements are inserted in the electro-acoustic scheme described by Alcoverro and Le Pichon (2004) between manifold and pipe or between the radiating impedance and the pipes at the input of the circuit. The capillaries could be inserted elsewhere inside the model if required. The frequency response is calculated from this modified electro-acoustic scheme by a matrix method. Using this method, noise reduction could also be simulated but is not presented here. These results are summarized in Fig. 5. The simulations correctly predict both the position of the resonance peaks and the amplitude for each of the systems tested in the field. The system with no acoustic resistance at the inlets is predicted to yield resonance at 2.65 Hz, 15 dB above the system with intermediate resistance. The system with 4.24 MΩ capillaries shows an 11-dB attenuation near 3 Hz. This is exactly what was observed in the recorded data (Fig. 4). The anti-aliasing filter applied to the recorded data complicates comparison of the higher frequency resonance peak at 7.95 Hz. Only the synthetics calculated for the system with capillaries at the inlets with acoustic resistance

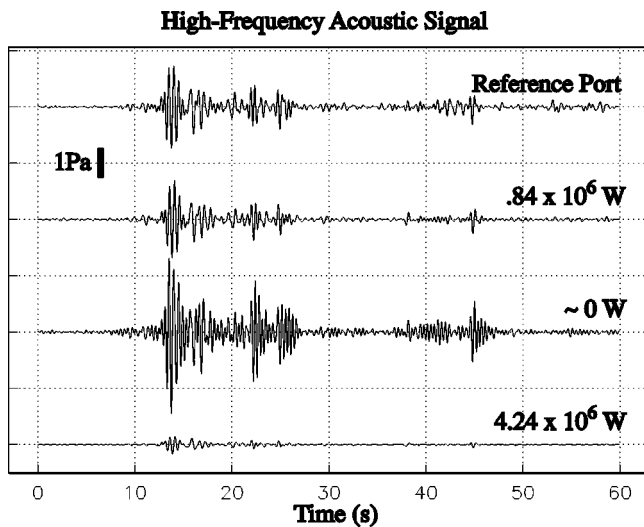


FIG. 6. An acoustic signal recorded by the reference site and the eight-arm filters. As expected, the system with the high acoustic impedance capillaries at the inlets overattenuates both the noise and the signal. The system with no acoustic resistance at the inlets (second trace from the bottom) produces highly resonant recordings. Only the system with the intermediate acoustic resistance capillaries ($0.84 \text{ M}\Omega$) installed at the inlets provided a recording that closely resembles the one provided by the reference site.

equal to the characteristic resistance of the pipes leading to the primary summing manifold are predicted to have no resonance peaks.

B. Analysis of a signal

On 21 November, 2001, a large signal was recorded by all systems in this experiment. Analysis of IS57 array data places the source at an azimuth of 240° . The phase velocity of the energy was 330 m/s . Although the cause of this signal remains unknown, the signal was coherent across the IS57 array and the simple spatial filters considered in this experiment, and provides an opportunity to examine the effect of capillaries on incident signals. As shown in Fig. 6, the incident wavetrain spans $\sim 40 \text{ s}$. The recordings of the signal via the reference site, and via the simple filter equipped with low impedance inlets, are similar, however the data from the filter exhibits a substantial coda. The signals recorded by the reference system are more impulsive. The system with intermediate acoustic resistances ($1.202 \text{ M}\Omega$) installed at the inlets produced a recording of the event that matches, much more closely, the character of the reference recording. The system with high resistance capillaries ($4.3 \text{ M}\Omega$) installed at the inlets produced a heavily attenuated copy of the signal. As predicted by the theory, the resistance above the characteristic impedance of the pipe overattenuates signal and noise at frequencies above 0.1 Hz by 11 dB . The high-impedance capillaries are predicted to low-pass filter the data below 0.2 Hz . The high inlet resistance that has proved able to remove much of the resonance has had an obvious, negative, effect on the incident signal.

V. TESTS OF CAPILLARIES ON ROSETTE FILTERS

The theory of Alcoverro and LePichon (2004) has been used successfully to estimate the effect of acoustic resistance

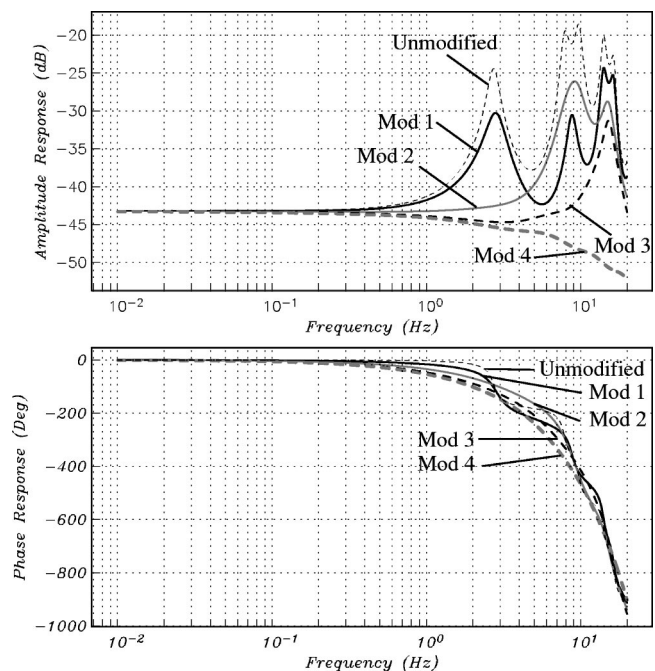


FIG. 7. Simulations of the original and modified 70-m aperture rosette filter frequency responses are shown in this figure. The systems with capillaries at the inlets (modification 1), at secondary summing manifolds in the pipes leading to the primary summing manifold (modification 2), at both the inlets and the secondary summing manifolds (modification 3), at the inlets and both summing manifolds (modification 4) are represented by the heavy black, heavy gray, heavy dashed black, and heavy dashed gray curves, respectively. The unmodified system is represented by the light dashed curves.

at the inlets of simplified rosette filters on resonance. The theory has further indicated that adding acoustic resistance at the inlet equal to the characteristic impedance of the 27-m -long pipe will remove all evidence of resonance near 2.65 Hz . Having validated the theory on simple filters, we now turn to full rosette filters.

A. Predictions of noise reduction in 70-m rosettes

We begin with the 70-m rosette filters in use at the IS57 infrasound array and shown in detail in Fig. 2. As shown in Fig. 7 (heavy solid black curve; modification 1), adding $2.07 \text{ M}\Omega$ acoustic resistance at the inlets, equal to the characteristic impedance of the 8-m pipe connecting the inlets with the secondary summing manifolds, reduces the resonance peak corresponding to the first fundamental frequency inside the 8-m pipes, near the Nyquist frequency. As shown by the heavy solid gray curve in the same figure, installing $1.202 \text{ M}\Omega$ acoustic resistance (capillaries 0.051 and 0.024 m in diameter) at each secondary summing inside the 27-m pipes leading to the primary summing manifold (modification 2) removes the main resonance peak near 2.65 Hz and its odd overtone near 7.95 Hz . The resistance at the secondary summing manifolds removes the fundamental resonance peak at 2.65 Hz and the first, odd, overtone near the Nyquist, but leaves the fundamental resonance occurring in the shorter pipes in the rosette clusters. The 180° phase shift that occurs at 2.65 Hz in the open system is blurred across a broader frequency band in the system that includes capillaries at the secondary summing manifolds.

It is possible to attack both the resonance occurring in

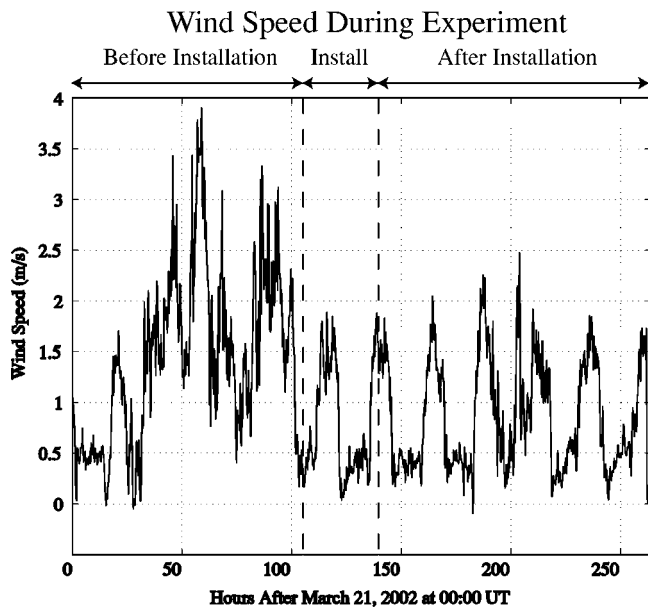


FIG. 8. Wind speeds during the experiment with the 70-m rosette filters.

the 27-m pipes, and the resonance in the 8-m pipes, by adding capillaries at the secondary summing manifolds and at the inlets (modification 3). A simulation of this is represented by the heavy dashed black curve in Fig. 7. Resonance energy above 5.0 Hz can be largely removed by modifying the filter in this manner. The resonance peak observed above 10 Hz (the heavy dashed black curve in Fig. 7) is due to resonance in the pipe that connects the primary summing manifold to the sensor. This can be removed by installing another 2.07 M Ω capillary adjacent to the primary summing manifold in that pipe. Synthesis of this modification (the fourth) is represented by the heavy dashed gray curves in Fig. 7.

As pointed out by Hedlin *et al.* (2003), the additional labor of installing capillaries at the 144 inlets and in the pipe between the primary summing manifold and the sensor is unlikely to be beneficial as the plane wave response of the filter to nonvertically incident signals severely attenuates energy above 3–4 Hz. The response of the system with capillaries at the primary and secondary summing manifolds as well as at the inlets is also not flat above 1 Hz. For these reasons, we proceeded with a test of capillaries at the secondary summing manifolds only.

B. Experiments with 70-m filters

To test the theoretical predictions, we installed 1.202 M Ω capillaries adjacent to the secondary summing manifolds in the main pipes leading to the primary summing manifold at the “L2” site in the array IS57. The other three 70-m filters (L1, L3, and L4 identified in Fig. 1) were left unchanged to provide a reference. We also collected pressure data with a microbarometer attached to a single inlet. The wind speed during this experiment before, during and after the modifications at L2 is shown in Fig. 8. The experiment spanned 11 days in March 2002. Diurnal variations in the wind speed are seen throughout the experiment with wind

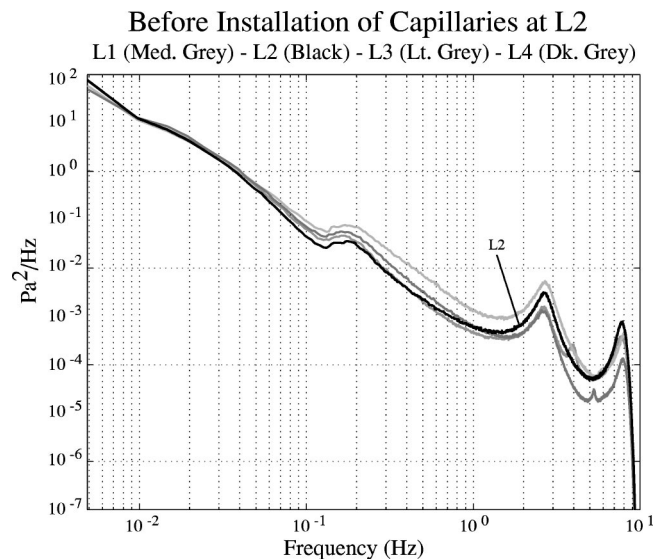


FIG. 9. Spectral density estimates taken from data collected at four 70-m rosette filters before any were modified. All spectra show significant resonance peaks at 2.65 and 7.95 Hz.

speeds varying from near zero to above 3.5 m/s. The wind speeds before the modification of L2 were slightly higher than those recorded after the change.

Stacked power spectral estimates from all 70-m filters prior to the installation of the capillaries at site L2 are shown in Fig. 9. The resonance peaks are evident in the data from all sites. In Fig. 10 we show a photograph of one capillary and a diagram showing its location in relation to the secondary summing manifold. The installation of each capillary required one person \sim 2 h. In Fig. 11, we show power spectral estimates taken from data collected after all eight capillaries were installed in the L2 filter. The results are striking. Very significant resonance peaks are observed in the data from the unmodified filters. No evidence of resonance in the data from the modified filter is seen below 5.0 Hz. The only sign of resonance, as predicted by the calculations shown in Fig. 7, is seen between 5 and 8 Hz. This is beyond the optimal pass-band of the 70-m rosette filters (Hedlin *et al.*, 2003).

C. Theory and experiment with 18-m rosette filters

As predicted in Fig. 3, resonance is expected to become apparent in data collected via the 18-m rosette filter at \sim 2.0 Hz. It is expected to rise to \sim 15 dB above background by 10 Hz. The problem with resonance in the 18-m rosette filters is less significant than it is in the large-aperture rosette filter considered in the previous section. As shown in Fig. 12, simulations of 18-m rosette filter frequency response equipped with impedance matching capillaries at both the secondary summing manifolds and at the inlets indicate that these can further improve the amplitude response of the filter. However, the phase shift is predicted to deviate from zero at frequencies below 1 Hz. In Fig. 12, we show the predicted amplitude and phase response of three 18-m rosette filters. The unmodified filter response is given by the dashed curve. The first resonance peak is centered just above 11 Hz and is due to resonance between the summing manifolds. A 180 $^\circ$ phase shift occurs at that frequency. A second reso-

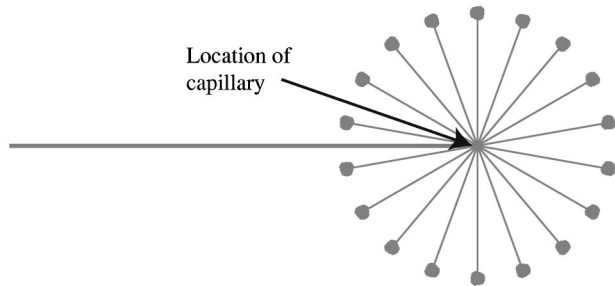
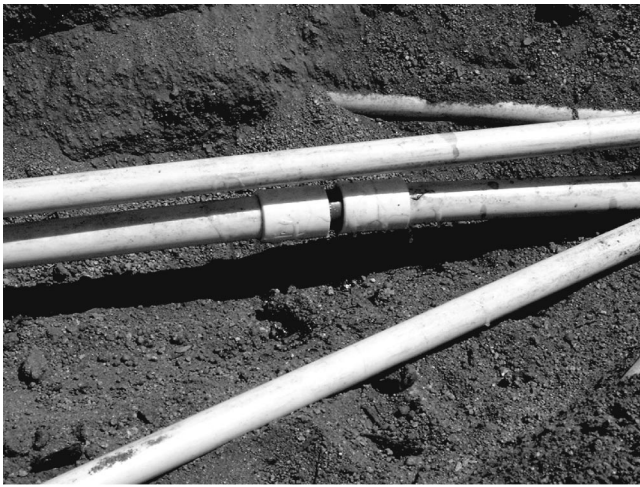


FIG. 10. A photograph of a capillary plug installed in a pipe connecting the primary and secondary summing manifolds. The capillary was installed 30 cm from the secondary summing manifold as shown in the lower diagram. The other pipes shown in the photograph connect the secondary summing manifold with inlets.

nance peak is observed near 19 Hz and is due to resonance in the inlet clusters between the inlets and the secondary summing manifolds. Installing 2.07 M Ω capillaries at the inlets removes the higher frequency resonance peak but leaves most of the resonant energy below the Nyquist frequency at 10 Hz. Installing 2.07 M Ω capillaries adjacent to the second-

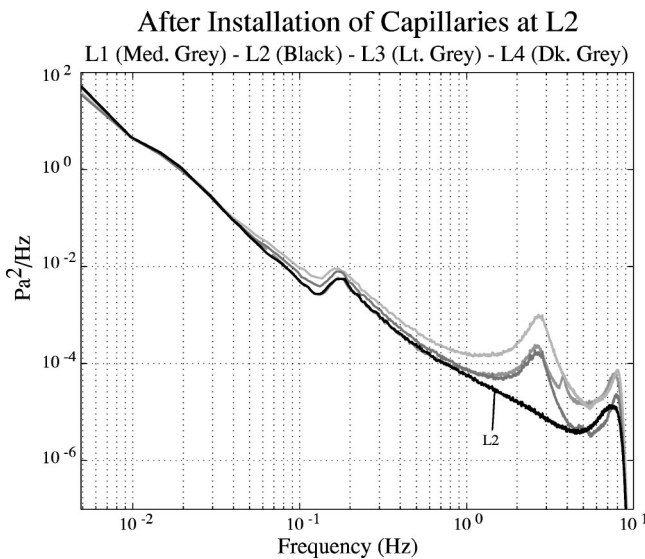


FIG. 11. Spectral density estimates taken from data collected after the capillary plugs were installed at site "L2" (black curve). The capillaries have removed the resonance peak at 2.65 Hz.

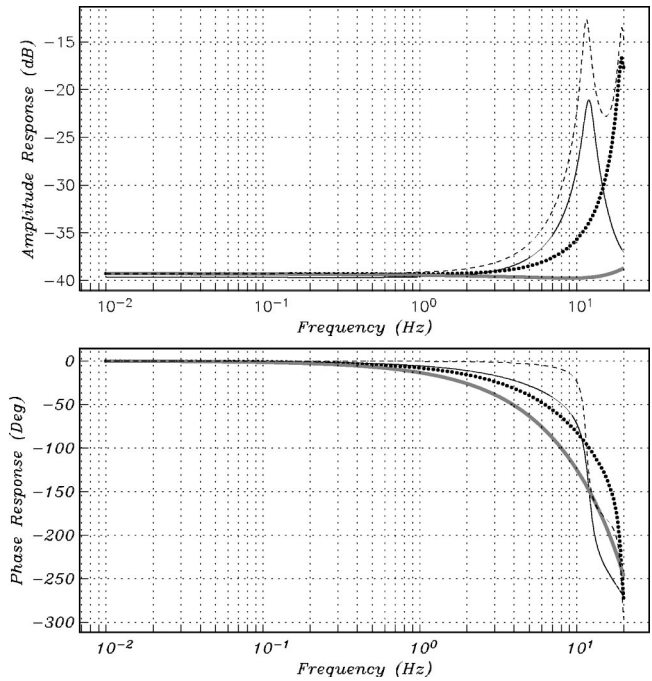


FIG. 12. Theoretical predictions of amplitude and phase response for the 18-m filter with and without capillaries. The unmodified filter is represented by the dashed curves. The filter with capillaries at the inlets, at the secondary summing manifolds, and at both locations are represented by the solid black, dotted black, and solid gray curves.

ary summing manifolds in the pipes leading to the primary summing manifold effectively removes the resonance between the summing manifolds. The amplitude response is improved by 13 dB at 10 Hz (dotted black curve in Fig. 12). When the system is equipped with capillaries at both inlets and secondary manifolds, the frequency response is flat up to 20 Hz (gray curve in Fig. 12). In Fig. 13, we show stacked spectra taken from all four 18-m rosette filters at the IS57 array before any were modified. The unmodified 18-m filters are inferior to the modified long-period 70-m filter at all frequencies. As shown by Hedlin *et al.* (2003) and by Alcoverro and LePichon (2004), unmodified 70-m filters outperform the 18-m filters only at frequencies below ~ 1.0 Hz.

Figure 14 shows spectral density estimates from the reference site and from an 18-m rosette filter before and after the capillaries were installed at the secondary summing manifolds. The capillaries remove the resonance above 1 Hz. The overall decrease of noise levels at all frequencies above 0.05 Hz is due to atmospheric conditions and is not due to the capillaries. Data from the unmodified 18-m filters showed the same decrease in energy.

VI. CONCLUDING REMARKS

Our research indicates that the reflections that produce resonance below 10 Hz in the 18- and 70-m rosette filters occur at the summing manifolds connected to the end of the longer pipe. Reflections that occur at the inlets artificially inflate spectral amplitudes at frequencies above 5 Hz. We find the theoretical method of Alcoverro and LePichon (2004) accurately predicts the frequency response of rosette

After Installation of Capillaries at L2

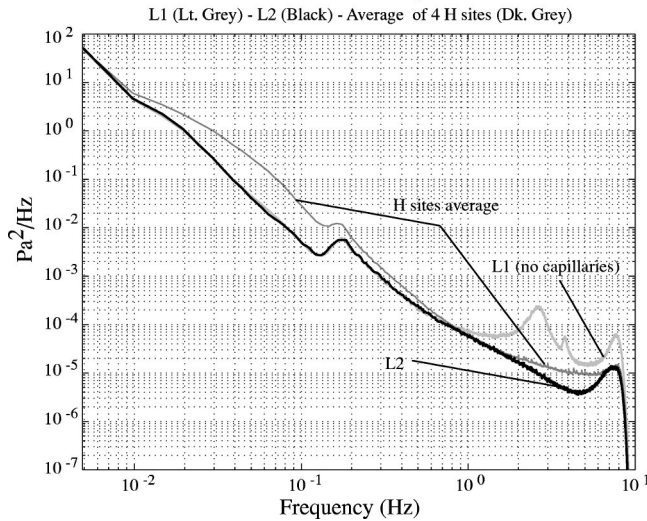


FIG. 13. Spectral density estimates from sites L1, L2 (both 70-m rosettes) and the 18-m rosettes. The data were taken after the L2 filter (black curve) was modified with the capillaries but before any modifications were made to the 18-m filters. The modified L2 70-m filter outperforms the unmodified 18-m filters at all frequencies. The resonance energy is first seen in the data from the 18-m rosette filters (the dark gray curve) at ~ 1 –2 Hz. The peak observed in all spectra between 0.1 and 0.2 Hz is due to microbaroms (Donn and Posmentier, 1967).

filters. The method also accurately predicts the effect of capillaries installed at any point inside the filters.

We conclude that the greatest gain in performance of the rosette filters is produced by installing impedance matching capillaries adjacent to the secondary summing manifolds in the pipes leading to the primary summing manifold. Installation of capillaries at the inlets is physically demanding, due to the large number of inlets involved, and offers little if any benefit to sensitivity of the filters due to the plane-wave response of the filters. As shown by Hedlin *et al.* (2003), the

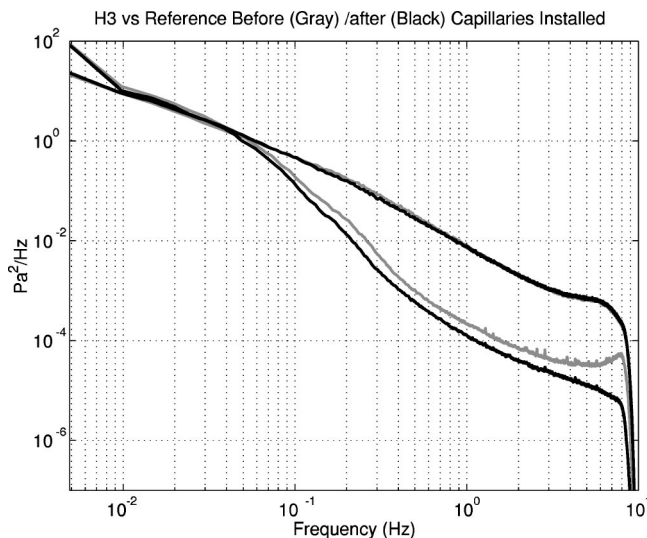


FIG. 14. Spectral density estimates from the reference site (upper curve) and from an 18-m rosette filter before (gray curve) and after (lower black curve) the capillaries were installed at the secondary summing manifolds. The capillaries have removed the resonance above 1 Hz. The observed decrease of noise levels at all frequencies above 0.05 Hz is not due to the capillaries but is due to atmospheric conditions.

plane wave response of the 70-m system is strongly dependant on frequency above 3 Hz. Above 5.0 Hz, much of the energy is attenuated. Adding capillaries at the inlets attacks resonant energy in a band that is already of limited value for the study of infrasonic signals.

There is an additional, practical, motive for not installing capillaries at the inlets but placing them in the interior of the filter. The capillary plugs installed in the 70- and 18-m rosette filters have an inner diameter of 2.4 mm. The desired acoustic resistance was achieved by varying the length of the capillary. Although the capillaries have been shown to provide an effective reduction of resonance inside the rosette filters, it will take long-term exposure to the elements to determine if the capillaries will, in practice, remain clear and remain effective without adverse side effects. If, with time, the capillaries become partially, or fully, occluded, the response of the overall rosette filter will be degraded. A fully occluded capillary will reduce the overall reduction of long-period noise provided by the filter by rendering useless an entire rosette of inlets. Partially occluded capillaries would have other adverse effects on the data as we have seen; the phase shift of the filter is determined, to some extent, by the capillary resistances. The rosette filters are not used alone, but as part of an array. It is essential that the amplitude and phase response of all filters in the array be matched. If changes in the acoustic resistance of the capillaries changes with time, and in a manner that differs between filters, the overall response of the array could be degraded. A bias could be introduced into the azimuths obtained from the array data.

Regardless of the location of the capillaries, theoretical tests using the method presented by Alcoverro and LePichon (2004) are required to determine the effect of one or more of the capillary plugs becoming partially or fully blocked by water, insects, etc. on the response of the individual filters and of the entire array. The long-term maintenance of these systems might involve periodic inspection and clearing of the capillary plugs.

Rosette spatial filters are used at almost all IMS infrasound arrays. All new filters include resonance-suppressing capillaries installed adjacent to the secondary summing manifolds, as described in this paper. Many of the existing filters are now being retro-fitted with capillaries. As a result, we anticipate an abundance of data in the near future will allow researchers to address these issues.

ACKNOWLEDGMENTS

The authors are indebted Chris Hayward (SMU) and Doug Christie (CTBTO) for suggesting that we experiment with capillaries. Frank Vernon, Jennifer Eakins, and Glen Offield provided the real-time data link. Clint Coon provided field assistance. Many helpful comments were provided by Lou Sutherland and two anonymous reviewers. Funding was provided by the Defense Threat Reduction Agency under Contract No. DTRA01-00-C-0085. Funding for the rosette filters used in this study was provided by the Defense Threat Reduction Agency (under Contract No. DTRA01-00-C-0085), the Provisional Technical Secretariat (PTS) of the UN

Comprehensive Test Ban Treaty Office in Vienna, and the US Army Space and Missile Defense Command (SMDC) University Research Initiative (URI).

Alcoverro, B. (1998). "Proposition d'un systeme de filtrage acoustique pour une station infrason IMS," CEA-DASE Scientific Report No. 241.
Alcoverro, B., and Le Pichon, A. (2004). "Design and optimization of a noise reduction system for Infrasound measurements using elements with low acoustic impedance," J. Acoust. Soc. Am.
Burrige, R. (1971). "The acoustics of pipe arrays," Geophys. J. R. Astron. Soc. **26**, 53–70.
Christie, D. R. (1999). "Wind-Noise-Reducing-Pipe Arrays," Report IMS-

IM-1999-1, Comprehensive Nuclear-Test-Ban Treaty Organization, Vienna Austria.
Christie, D. R., Vivas Veloso, J. A., Campus, P., Bell, M., Hoffmann, T., Langlois, A., Martysevich, P., Demirovich, E., and Carvalho, J. (2001). "Detection of atmospheric nuclear explosions: the infrasound component of the International Monitoring System," Kerntechnik **66**, 96–101.
Daniels, F. B. (1959). "Noise reducing line microphone for frequencies below 1 c/s," J. Acoust. Soc. Am. **31**, 529.
Donn, W. L., and Posmentier, E. S. (1967). "Exploring the Atmosphere with Nuclear Explosions," Rev. Geophys. **5**, 53.
Grover, F. H. (1971). "Experimental noise reducers for an active microbarograph array," J. Geophys. Soc. **26**, 41–52.
Hedlin, M. A. H., Alcoverro, B., and D'Spain, G. (2003). "Evaluation of rosette infrasonic noise-reducing spatial filters," J. Acoust. Soc. Am. **114**, 1807–1820.

Simultaneous acoustic channel measurement via maximal-length-related sequences^{a)}

Ning Xiang^{b)}

School of Architecture, and Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, New York 12180

John N. Daigle

Department of Electrical Engineering, Anderson Hall, University of Mississippi, University, Mississippi 38677

Mendel Kleiner

School of Architecture and Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, New York 12180

(Received 8 October 2004; revised 13 January 2005; accepted 17 January 2005)

A wide variety of acoustic systems has multiple sources and receivers. This paper proposes a technique for making acoustic measurements simultaneously for multiple sources. The proposed technique features a collection of excitation signals of the maximum-length sequence (MLS) and MLS-related classes. Each signal in the set has a pulse-like autocorrelation function, and the cross-correlation functions between arbitrary pairs of signals drawn from the set have peak values that are significantly lower than the peak magnitude of the autocorrelation functions. The proposed method is particularly valuable when characterization of multisource, multireceiver system must be accomplished in a limited time period. Both simulation and experimental results are presented that demonstrate the feasibility and fidelity of the proposed techniques in characterizing acoustic systems. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1868252]

PACS numbers: 43.28.Vd, 43.28.Tc, 43.58.Gn [EJS]

Pages: 1889–1894

I. INTRODUCTION

A previous paper¹ presented a technique using reciprocal pairs of maximal-length sequences (MLSs) in simultaneous dual-source channel measurements. Impulse responses between two separate sources and one or several receivers of acoustic systems can be determined simultaneously. The simultaneous dual-source measurements exploit excellent cross-correlation properties of reciprocal MLS pairs. Once the number of simultaneous sources is higher than two, straightforward use of reciprocal MLS pairs cannot meet the need. Special attention must then be given to generation of a larger set of source signals that has acceptable properties for performing the simultaneous measurements.

Because of their favorable cross-correlation properties, certain sets of MLSs and MLS-related sequences have gained considerable attention in spread spectrum communication systems.² They are, however, not yet widely applied to acoustic systems, particularly in the area of simultaneous acoustic measurements. Recent research in sound propagation through atmosphere^{3,4} particularly calls for the simultaneous measurement technique. For this purpose, we propose here the application of a set of the MLSs and MLS-related classes in which all sequences of the set possess a *pulse-like* periodic autocorrelation function, while the periodic cross-correlation function between any pair of sequences drawn from the set has a peak value that is significantly lower than the peak value of the autocorrelation function. In general, the proposed technique is particularly valuable where multi-

source, multireceiver system characterization tasks must be accomplished in a limited time period.

In Sec. II we discuss the desired correlation properties of acoustic excitation signals that would be needed to accomplish simultaneous multiple source measurements in an ideal case. We then briefly describe suitable sets of acoustic signals based on binary MLSs and MLS-related sequences. This briefing relies heavily on the literature in the spread spectrum communication.^{2,5,6} Using these sets, simultaneous channel characterization can be accomplished to a suitable degree of accuracy. In Sec. III we present simulation and field measurements that demonstrate the feasibility of simultaneous channel characterization. Finally, we draw some conclusions in Sec. IV.

II. CORRELATION PROPERTIES OF MLS AND RELATED SEQUENCES

For simultaneous multiple acoustic source measurements (SMASM), let n and p denote the number of sources and receivers, respectively. Then, the p -element vector of received signals, \mathbf{r} , can be obtained from

$$\mathbf{r} = \mathbf{h} * \mathbf{s}, \quad (1)$$

where \mathbf{s} is the $n \times 1$ source signal vector, $\mathbf{h} = [h_{ij}]$ is the impulse response matrix of dimension $p \times n$, and “*” denotes the matrix periodic (circular) convolution operator. Each element of \mathbf{r} , \mathbf{h} , and \mathbf{s} is, again, a discrete function of time, but time dependency is suppressed for brevity. If the periodic cross-correlation function (PCCF) of excitation signals is

^{a)}Parts of this work have been presented at the 18th ICA, Kyoto, Japan.

^{b)}Electronic mail: xiangn@rpi.edu

$$\mathbf{s} \otimes \mathbf{s}^T = [\delta], \quad (2)$$

where the operator \otimes stands for the cross correlation, $[\delta]$ denotes the matrix each of whose diagonal elements is a discrete periodic delta function and each of whose off-diagonal elements are the all-zero function, then the impulse response matrix can be directly determined from¹

$$\mathbf{h} = \mathbf{r} \otimes \mathbf{s}^T. \quad (3)$$

Equation (3) indicates that the individual impulse response sequences, h_{ij} , between source j and receiver i , can be conveniently determined by cross correlating each received signal with each source signal provided the source signals possess the desired correlation property expressed in Eq. (2). Lüke⁵ pointed out, however, that signals that exactly fulfill this condition [Eq. (2)] neither exist nor can be constructed. Nevertheless, a careful selection or combination among MLSs of the same degree may yield a set of sequences approximately possessing the desired correlation property.

An n -stage linear feedback shift-register device can generate a binary MLS, $\{a(i)\}$, with a period of $L = 2^n - 1$. The positive integer, n , is said to be the degree of the MLS.^{7,8} The sequence obtained through correlation of one bipolar MLS $\{m(i)\}$ with $m_i = 1 - 2a_i$ or MLS-related sequence with another of the same degree, or cross correlation, is of practical significance for the SMASM technique, which has been known in the communication community for many years. They are, however, considerably less widely known in the acoustics community; therefore, a brief description of cross-correlation properties of various sets of sequences drawn from MLS and MLS-related classes is sufficient in the following.

A. Preferred MLS pairs and maximum connected sets

The sequence $\{b(i)\}$, where $b(i) = a(di)$ and d is appropriately chosen, is itself an MLS, known as a *factor-d decimation of the sequence* $\{a(i)\}$,⁶ the cross correlation of the sequences $\{b(i)\}$ and $\{a(i)\}$, will yield very small values. If the degree of an MLS is a multiple of 4, then a decimation factor⁹ $d(n) = 2^{(n+2)/2} - 1$ will lead to an MLS pair whose PCCF takes on only four values. For such decimations, the upper bounds on the four-valued PCCF is

$$l_4(n) = \frac{d(n)}{2^n - 1}. \quad (4)$$

If the degree of an MLS is not a multiple of 4, then an appropriate decimation of that MLS will yield pairs of MLSs having relatively small three-valued PCCF.^{6,10} For MLSs whose degrees are not a multiple of 4, some of decimation factors are of the form $d = 2^k + 1$ or $d = 2^{2k} - 2^k + 1$, where k is chosen such that $n/\text{gcd}(n,k)$ is odd,⁶ with $\text{gcd}()$ standing for the greatest common divider, the cross-correlation bounds take on preferred small values expressed by

$$l_3 = \frac{2^{\lfloor (n+2)/2 \rfloor + 1}}{2^n - 1}, \quad (5)$$

where $\lfloor x \rfloor$ denotes the integer part of the real number x . This paper will refer to both the three-valued and four-valued MLS pairs as preferred pairs. Their bound values are of prac-

TABLE I. Cross-correlation bound values of preferred and reciprocal MLS pairs, of Gold and Kasami sequences.

Degree	Period length	Reciprocal MLS pairs	Preferred MLS pairs & Gold sequences	Kasami sequences
8	255	0.1216	0.1216	0.067
9	511	0.0881	0.0646	
10	1023	0.0616	0.0635	0.0322
11	2 047	0.0437	0.0317	
12	4 095	0.0310	0.0310	0.0159
13	8 191	2.20E-2	1.57E-2	
14	16 383	1.56E-2	1.57E-2	7.87E-3
15	32 767	1.10E-2	7.84E-3	
16	65 535	7.80E-3	7.83E-3	3.92E-3
17	131 071	5.52E-3	3.91E-3	
18	262 143	3.90E-3	3.91E-3	1.96E-3
19	524 287	2.76E-3	1.95E-3	
20	1048 575	1.95E-3	1.95E-3	9.77E-4
21	2 097 151	1.38E-3	9.77E-4	
22	4 194 303	9.76E-4	9.77E-4	4.89E-4
23	8 388 607	6.90E-4	4.88E-4	
24	16 777 215	4.88E-4	4.88E-4	2.44E-4

tical significance for the SMASM techniques.

In addition to the preferred MLS pairs, reciprocal pairs of every binary MLS of degree n also possess small-valued PCCF (Refs. 1, 11) bounded by

$$l_r = \frac{2^{(n+2)/2} - 1}{2^n - 1}, \quad (6)$$

but not limited to three or four values.

Table I lists all bound values of preferred and reciprocal MLS pairs between degree 8 and 24 for an easy comparison. Figure 1 also shows the smallest bounds (logarithmic peak values) of the PCCF of preferred and reciprocal MLS pairs between degree 8 and 24. A careful comparison between columns in Table I and between Eq. (5) and Eq. (6) reveals that the small-valued PCCFs of the preferred MLS pairs are close to that of reciprocal ones for even-numbered degrees, while for the odd-numbered degrees the preferred MLSs are ap-

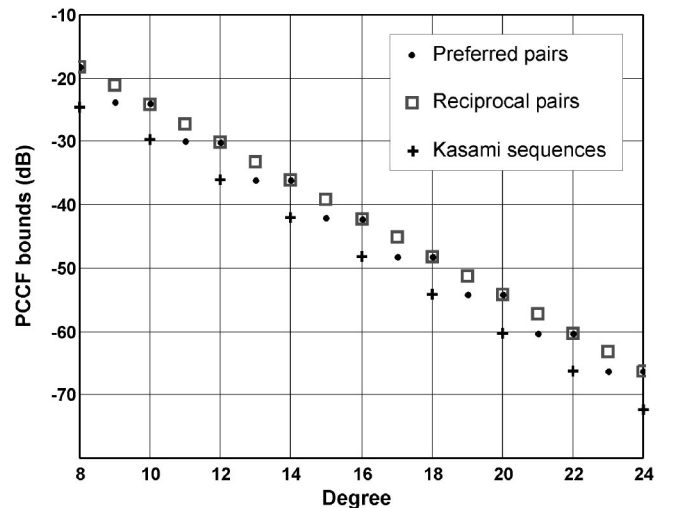


FIG. 1. Bound values of the cross-correlation functions of preferred and reciprocal MLS pairs along with bound values of Kasami sequences. The bound values are expressed in dB relative to the peak value of their auto-correlation function.

proximately one-half that for the reciprocal MLSs. For the current applications, one should not necessarily insist on having three- or four-valued PCCF as long as both their PACF and PCCF approximate the desired condition expressed in Eq. (2).

When specific applications require up to four simultaneous source signals, however, even one preferred MLS pair along with their individual reciprocal pairs will not yield the preferred smallest bound values among four of them. To be more precise, let $\{a(i)\}$ and $\{b(i)\}$ denote a preferred MLS pair of degree 14, and $\{a(i)\}$ and $\{a_r(i)\}$, $\{b(i)\}$, and $\{b_r(i)\}$ denote reciprocal pairs of each individual MLS, respectively. The PCCFs between $\{a(i)\}$ and $\{b(i)\}$, and between $\{a_r(i)\}$ and $\{b_r(i)\}$ are bounded by a preferred value of 0.0157, while the PCCFs between $\{a(i)\}$ and $\{a_r(i)\}$, and between $\{b(i)\}$ and $\{b_r(i)\}$ are bounded at a preferred value of 0.0156, as listed in Table I. However, the PCCFs between $\{a(i)\}$ and $\{b_r(i)\}$, and between $\{b(i)\}$ and $\{a_r(i)\}$ are bounded by 0.03, which is roughly twice the preferred small values.

Practical applications often need more than two simultaneous source channels. While a pair of MLSs selected from two different preferred pairs of the same degree may not possess the preferred small-valued PCCF as listed in Table I and plotted in Fig. 1, some pairs selected from different pairs do. A set of the MLSs for which each pair of the set has this preferred small PCCF values is referred to as a connect set. The largest possible connected set is termed a maximum connect set (MCS) (Ref. 10) and the size of such a MCS is denoted by M_n . In the literature,⁶ one can find the values of M_n for $n \leq 16$ with $M_n \leq 4$. The length of MLSs of these degrees and their reasonably low PCCF bound values are of practical interest in a wide variety of acoustic measurements. However, the number of MLSs in the MCS are strongly limited. One needs to use properly combined MLSs to form Gold and Kasami sequences.

B. Combination of MLS pairs

A binary Gold sequence $\{G_\rho(i)\}$ can be generated by combining a preferred, or a reciprocal binary MLS pair $[\{a(i)\}, \{b(i)\}]$ as¹⁰ $G_\tau(i) = a(i) \oplus b(i + \tau)$, with \oplus denoting addition module 2. In stepping τ point-by-point, a large number of Gold sequences result including the preferred MLS pair to yield a finite set $\{\{a\}, \{b\}, \{G_0\}, \{G_1\}, \dots\}$ containing exactly $L + 2$ sequences. Sequences in this finite set approximately fulfill the condition expressed in Eq. (2). PACFs of bipolar Gold sequences associated with $\{G_\tau\}$ are peaked at zero lag, but they also possess small-valued sidelobes. The PCCFs of Gold sequences possess the same bound value as that of the sidelobes of the corresponding PACF. In fact, the bound values are the same as those of the PCCF of the preferred and reciprocal MLS pairs from which the Gold sequences are derived.

For even-numbered degree n , a decimation from an MLS $\{a(i)\}$ with factor $\lambda(n) = 2^{n/2} + 1$ can lead to a pair $\{a(i)\}$ and $\{c(i)\}$ with $c(i) = a(\lambda i)$. In this case, $\{c(i)\}$ is not an MLS of degree n . A binary Kasami sequence $K_\tau(i)$ can be generated by combining $\{a(i)\}$ and $\{c(i)\}$ as $K_\tau(i) = a(i) \oplus c(i + \tau)$. In similar fashion, a set of Kasami se-

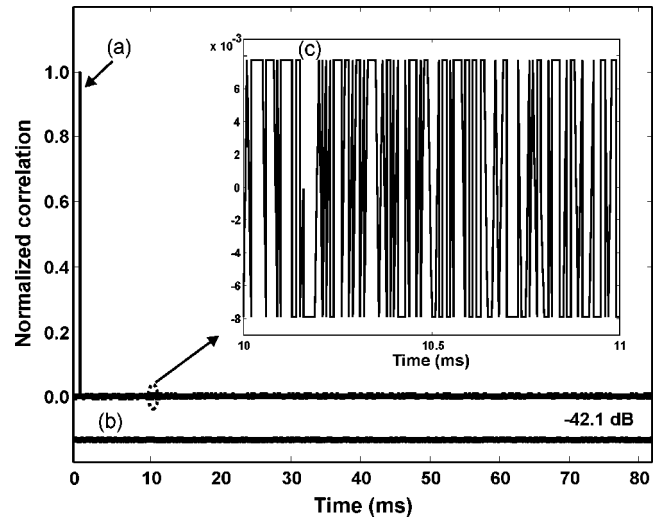


FIG. 2. Normalized correlation functions of Kasami sequences of degree 14. (a) Periodic autocorrelation function (PACF) of one sequence. (b) Periodic cross-correlation function (PCCF) between two Kasami sequences (shifted downwards beneath the autocorrelation function for a convenient comparison). The peak value of the PCCF is 42.1 dB lower than that of PACF. (c) Zoomed presentation of a segment from (a) showing three-valued sidelobes of the PACF. Their peak values are the same as those of the PCCF in (b).

quences $\{\{a\}, \{K_0\}, \{K_1\}, \dots\}$ can be constructed¹² containing $2^{n/2}$ sequences. The sidelobes of PACFs and the amplitude of PCCFs among these sequences present even lower bounds, approximately half of those of preferred pairs and Gold sequences

$$l_k(n) = \frac{\lambda(n)}{2^n - 1}, \quad (7)$$

where n must be a positive even number. Figure 1 also shows the bounds values of Kasami sequences in logarithmic scale along with those of preferred and reciprocal MLS pairs for comparison. The power spectral density functions of Gold and Kasami sequences are the same as that of MLSs, being of broadband nature and covering the entire frequency range. In order to demonstrate the excellent correlation properties of Gold and Kasami sequences, Fig. 2 shows the PACF and PCCF of a pair of Kasami sequences, derived from an MLS of degree 14 decimated by a factor of 129. In the example, the PCCF values and the sidelobes of the PACF are bounded by 0.007 87, or 42.1 dB below the peak value of the individual PACF. Unlike MLSs, the PACFs of Gold and Kasami sequences are pulse-like functions with small-valued sidelobes. They approximate the condition expressed in Eq. (2) to the required degree of accuracy, and they are equivalent to the MLSs drawn from MCSs with regard to their usefulness in the SMASM technique.

III. SIMULATION AND EXPERIMENTAL RESULTS

In this section we present simulation and experimental results to demonstrate usefulness of the lack of correlation among the MLSs from MCS and combined MLSs. Unlike applications in the spread spectrum technology, where MLS-related sequences are primarily used for coding or modulation, the acoustic measurement technique discussed in this paper employs the sequences directly for acoustic excita-

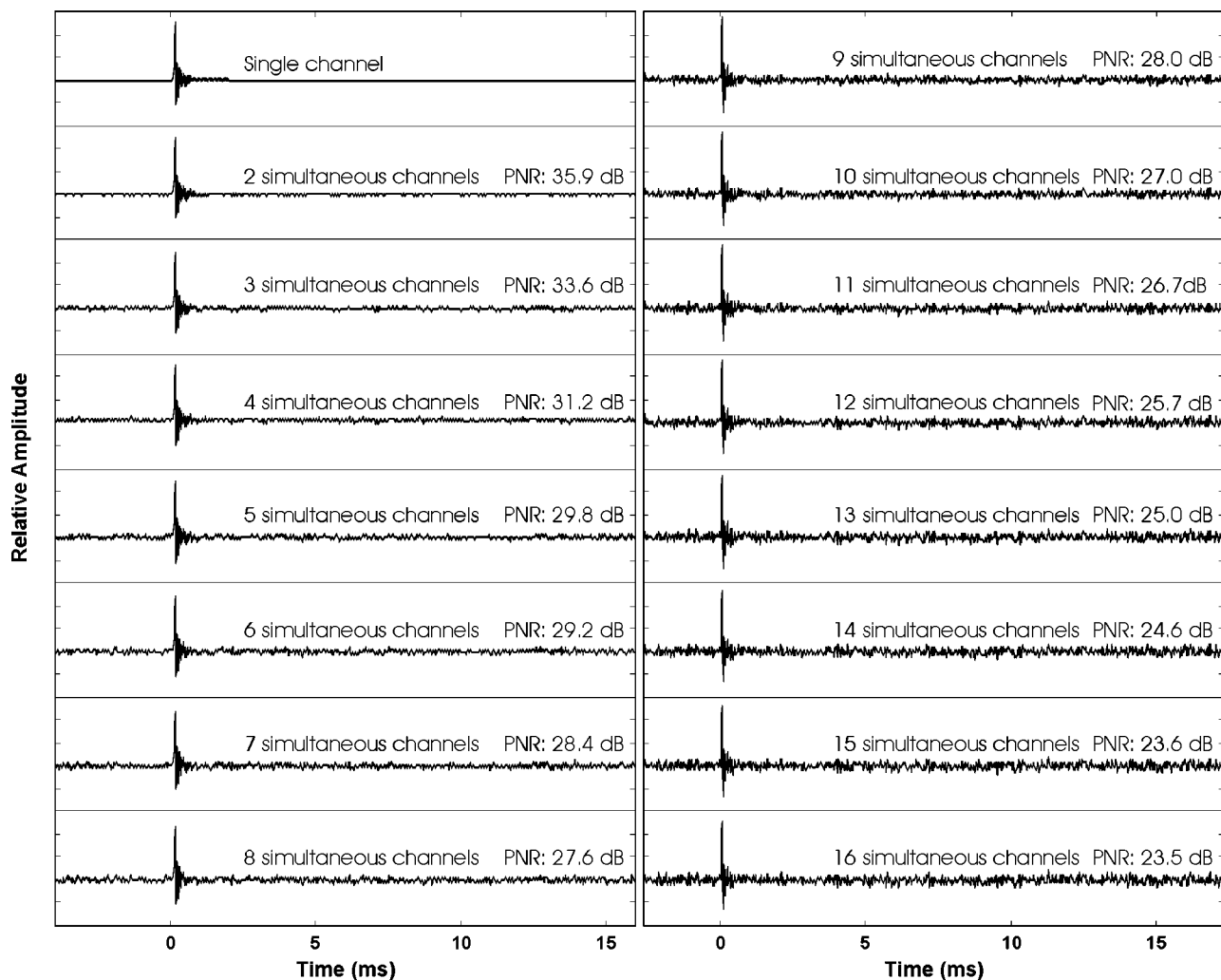


FIG. 3. Simulation results using digital filters to study the noise behavior of simultaneous multiple source measurements, Single channel, 2, 3,... and 16 simultaneous source channels are simulated. Impulse responses obtained using increasing number of simultaneous Kasami sequences of degree 14 are label by the number of simultaneous source channels and the peak-to-noise ratio (PNR).

tions. The acoustic channel responses are then cross correlated with the excitation sequences. Since these sequences are of length $2^n - 1$, conventional FFT cannot be employed directly. Fortunately the fast MLS transform (based on fast Hadamard transform) can be employed for MLSs and reciprocal MLSs.¹ When using combined sequences, however, one needs to exploit a specialized algorithm¹³ for calculating the PCCF.

A. Digital simulation

In order to understand the signal-to-noise ratio achievable in the SMASM applications, we apply a set of 16 bipolar Kasami sequences $\{\{k_1\}, \{k_2\}, \dots, \{k_{16}\}\}$ transmitted at a rate of 50 kHz to simulate acoustic channel experiments. Among them, $\{k_1\}$ is a single MLS of degree 14 and $\{k_2\}, \dots, \{k_{16}\}$ are Kasami sequences constructed by combining $\{k_1\}$ and $\{b\}$, with $b(i) = k_1(129i)$. The PCCFs among the 16 sequences possess the same small values as given in Table I.

Eighth-order low-pass Chebyshev filters, each having a cutoff frequency of 16 kHz, are used to simulate individual acoustic channels. Each individual sequence is filtered by

one low-pass filter separately. First one, then two, ..., then all 16 of the filter responses are summed to simulate the signal that would be received at one point and at a common point from two, three, ..., and all 16 simultaneous sources. The summed signals are then cross correlated with the MLS $\{k_1\}$ to obtain the impulse response associated with the source sending $\{k_1\}$. Figure 3 shows the impulse responses of a single source channel and of simultaneous multiple channels.

To quantify the quality of the impulse response, we define peak-to-noise ratio (PNR) as the ratio of its peak value to the rms value of noise, which is the difference between the exact value of the impulse response and the value calculated using our technique at each time epoch. More precisely, the signal received at receiver i is

$$r_i = \sum_{j=1}^{16} h_{ij} * s_j. \quad (8)$$

If we now correlate the received signal with s_k , we obtain the following approximation for h_{ik} :

$$\hat{h}_{ik} = h_{ik} + h_{ik} * (s_k \otimes s_k - \delta) + \sum_{j=1, j \neq k}^{16} h_{ij} * (s_j \otimes s_k). \quad (9)$$

Thus, the sequence

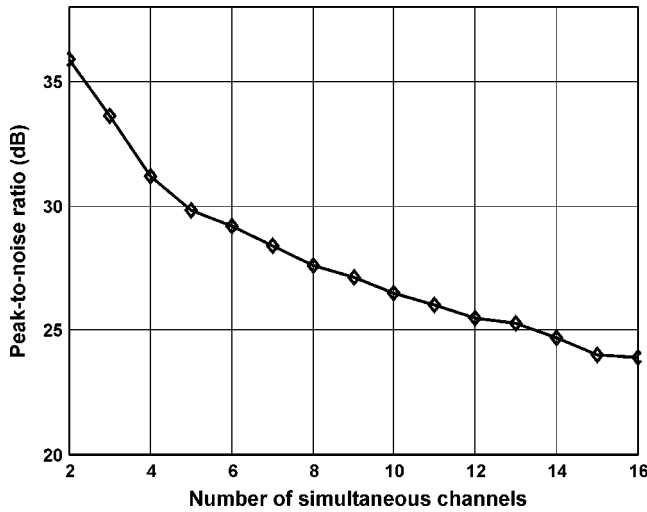


FIG. 4. Peak-to-noise ratios (PNR) as a function of simultaneous source channels, achieved using digital filters to simulate the acoustic channels. PNR obtained using Kasami sequences of degree 14 ranges from 35.9 dB down to 23.5 dB when the number of simultaneous source changes from 2 to 16.

$$e_{ik} = \hat{h}_{ik} - h_{ik} = h_{ik} * (s_k \otimes s_k - \delta) + \sum_{j=1, j \neq k}^{16} h_{ij} * (s_j \otimes s_k), \quad (10)$$

represents the error between the approximation for h_{ik} and its actual value, that is, it is a noise sequence. Thus

$$\overline{E_{ik}^2} = \frac{1}{L} \sum_{z=0}^{L-1} e_{ik}^2(x), \quad (11)$$

is the mean-squared error between h_{ik} and its approximation, \hat{h}_{ik} . If the autocorrelation function of s_k were a δ function, and s_j and s_k were uncorrelated for $j \neq k$, then the mean-squared error would be zero. Otherwise, the ratio of the peak value of h_{ik} to the rms noise value gives a measure of the quality of the estimate of h_{ik} .

Figure 4 shows the PNR achieved at individual numbers of simultaneous sources. The simulation results as shown in Fig. 3 and Fig. 4 indicate that the PNR in the SMASM decreases with increasing number of simultaneous source signals. The noise stems from the sidelobes of the autocorrelation and nonzero residuals of the cross-correlation operations as shown in Fig. 2. The higher the degree of the sequences used, the higher the PNR will be, as Fig. 1 implies. In a similar fashion, when cross correlating the resulting receiver signal in the simultaneous multiple source mode with MLS $\{k_2\}, \{k_3\}, \dots$, the impulse responses $\{h_{12}\}, \{h_{13}\}, \dots$ would be resolved.

The fast M-sequence transform¹ is used in an efficient calculation of cross correlation between MLS $\{k_1\}$ and the summed filter responses. When resolving impulse responses associated with combined MLSs $\{k_2\}, \{k_3\}, \dots$, the fast MLS transform cannot be used, but a specialized algorithm may be used for calculating PCCF.¹³

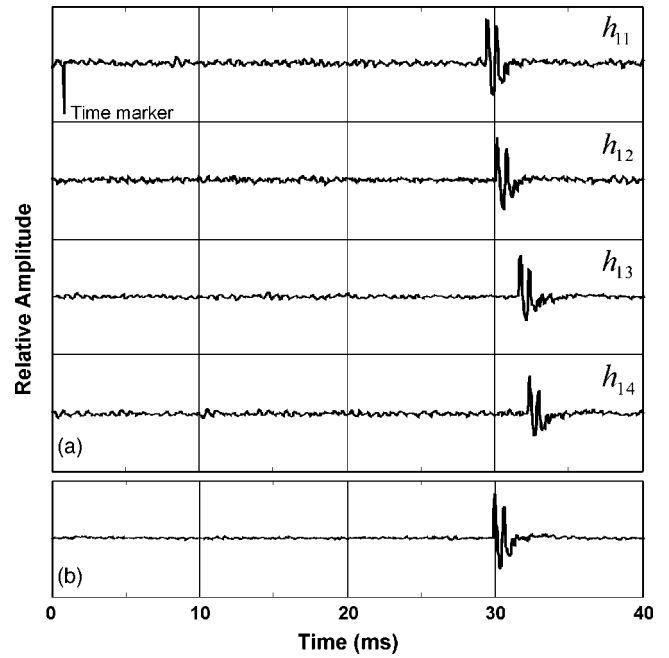


FIG. 5. Experimental results of times of flight for outdoor sound propagation analysis. (a) Four simultaneous sound sources and a single receiver. (b) Impulse response between a single source (#2) and the receiver (for comparison with h_{12}).

B. Experimental results

We now present a group of field measurement results achieved from exploratory experiments conducted outdoors. Four loudspeakers, which serve as sources, are separated 2 m from each other along a straight line. Perpendicular to this line, a microphone (receiver) is set 10 m away from the first sound source. The receiver to the farthest source is about 11 m. The sound sources and receivers are set 1.1 m above the outdoor ground surface. The sampling frequency for these measurements is 25 kHz. Four MLSs' selected from the MCS of degree 13 are used to simultaneously drive the sound sources. Figure 5(a) illustrates impulse response segments of the first 40 ms associated with four simultaneous sound sources. They are obtained by using the fast MLS transform¹ of the received receiver signal with each of the individual MLSs'. These simultaneous measurements of channel impulse responses yield slightly different arrival times due to different propagation distances of the acoustic wave, corresponding well to the geometry of the setup. Figure 5(b) shows the measured single source impulse responses of the channel between the second source and the receiver when only source two is energized. Although the single source measurement yields a better PNR, the single- and multiple source measurements clearly yield similar results.

An achievable PNR depends upon, among other things, available source power, the sensitivity and bandwidth of the sensors or receivers, the attenuation of the propagation channels, and the length of sequences. In practice, the achievable PNR can also be influenced by other unwanted noise, background noise, nonlinear or time-variant components within the system under test, resulting in a lower PNR. While SMASM cannot provide results of the identical quality that may be obtained through sequential measurement, a well-

designed SMASM can produce results meeting the requirements of all but the most demanding applications.

IV. CONCLUSIONS

The cross-correlation properties of maximum-length sequences (MLSs) and related signals make them good candidates for simultaneous multiple acoustic source measurements (SMASM). In this paper we have demonstrated that, by exploiting the cross-correlation properties of MLSs and MLS-related sequences—Gold and Kasami, in particular—SMASM becomes feasible. Both simulation and experimental results achieved from exploratory outdoor measurement have demonstrated the feasibility of using MLSs and combined MLSs for the SMASM technique. This proposed measurement technique would be especially useful when an application demands simultaneous, rather than sequential, excitation of multiple sources.^{3,4} When a sufficient number of single MLSs from a maximum connect set (MCS) can meet the need, a fast algorithm, the so-called fast MLS transform^{14–17} can be applied directly to retrieving impulse responses between simultaneous multiple sources and one or several receivers. Otherwise, additional Gold and Kasami sequences can be selected to meet the need. In the latter case, a specialized algorithm, such as the one documented in Ref. 13, must be applied.

ACKNOWLEDGMENTS

The authors thank Professor Dr. Hans D. Lüke for his valuable advice and inspiration. The authors also express their gratitude to Dr. H. Bass, Dr. R. Raspet, Dr. J. M. Sabatier, and Dr. R. R. Torres, who supported this work with enthusiasm. The authors would like to thank the associate editor and the reviewers for their insightful and constructive comments.

- ¹N. Xiang and M. R. Schroeder, "Reciprocal maximum-length sequence pairs for acoustic dual source measurements," *J. Acoust. Soc. Am.* **113**, 2754–2761 (2003).
- ²M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications Handbook* (McGraw-Hill, New York, 1994).
- ³D. K. Wilson, A. Ziemann, V. E. Ostashev, and A. G. Voronovich, "An overview of acoustic travel-time tomography in the atmosphere and its potential applications," *Acust. Acta Acust.* **87**, 721–730 (2001).
- ⁴A. Ziemann, K. Arnold, and A. Raabe, "Acoustic tomography as a method to identify small-scale land surface characteristics," *Acust. Acta Acust.* **87**, 731–737 (2001).
- ⁵H. D. Lüke, *Korrelationsignale* (Springer, Berlin, New York, 1992).
- ⁶D. V. Sarwate and M. B. Pursley, "Cross-correlation properties of pseudorandom and related sequences," *Proc. IEEE* **68**, 593–619 (1980).
- ⁷S. W. Golomb, *Shift Register Sequences* (Aegean Park, Laguna, CA, 1982).
- ⁸F. J. MacWilliams and N. J. Sloane, "Pseudo-random sequences and arrays," *Proc. IEEE* **64**, 1715–1729 (1976).
- ⁹Y. Niho, "Multi-Valued Cross-Correlation Functions Between Two Maximum Linear Recursive Sequences on Comma-Free Codes," Ph.D. dissertation, Dept. Elec. Eng., University of Southern California (1972).
- ¹⁰R. Gold, "Maximum recursive sequences with 3-values cross-correlation functions," *IEEE Trans. Inf. Theory* **IT-14**, 151–156 (1968).
- ¹¹T. A. Dowling and R. McEliece, "Cross-Correlation of Reverse Maximal-Length Shift Register Sequences," *SPL Space Programs Summary 37–53 III*, pp. 192–193 (1968).
- ¹²T. Kasami, "Weight distribution of Bose–Chaudhuri–Hocquenghem codes," in *Combinatorial Mathematics and Its Applications* (University of North Carolina Press, Chapel Hill, NC, 1969).
- ¹³J. N. Daigle and N. Xiang, "An FFT-Based Algorithm for Acoustic Measurements Using Coded Signals," *Proceedings of the International Congress on Acoustics IV*, pp. 2989–2992, 2004.
- ¹⁴M. Cohn and A. Lempel, "On fast M-sequences transforms," *IEEE Trans. Inf. Theory* **IT-23**, 135–137 (1977).
- ¹⁵A. Lempel, "Hadamard and M-sequence transforms are permutationally similar," *Appl. Opt.* **18**, 4064–4065 (1979).
- ¹⁶J. Borish and J. B. Angell, "An efficient algorithm measuring the impulse response using pseudorandom noise," *J. Audio Eng. Soc.* **31**, 478–488 (1983).
- ¹⁷N. Xiang, "Using M-sequences for determining the impulse responses of LTI-systems," *Signal Process.* **28**, 139–152 (1992).

Phase speed and attenuation in bubbly liquids inferred from impedance measurements near the individual bubble resonance frequency

Preston S. Wilson,^{a)} Ronald A. Roy, and William M. Carey

Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Room 101, Boston, Massachusetts 02215

(Received 30 July 2004; revised 21 December 2004; accepted 22 December 2004)

In the ocean, natural and artificial processes generate clouds of bubbles that scatter and attenuate sound. Measurements have shown that at the individual bubble resonance frequency, sound propagation in this medium is highly attenuated and dispersive. The existing theory to explain this behavior is deemed adequate away from resonance. However, due to excessive attenuation near resonance, little experimental data exists for a comparison with model predictions. An impedance tube was developed specifically for exploring this regime. The effective medium phase speed and attenuation were inferred from measurements of the surface impedance of a layer of bubbly liquid composed of air bubbles and distilled water, for void fractions from 6.2×10^{-5} to 5.4×10^{-4} and bubble sizes centered around 0.62 mm in radius. Improved measurement speed, accuracy, and precision is possible with the new instrument, and both instantaneous and time-averaged measurements were obtained. The phase speed and attenuation at resonance was observed to be sensitive to the bubble population statistics and agreed with an existing model [J. Acoust. Soc. Am. **85**, 732–746 (1989)], within the uncertainty of the bubble population parameters. Agreement between the model and the data reported here is better than for the data that was available when the model was originally published. © 2005 Acoustical Society of America.

[DOI: 10.1121/1.1859091]

PACS numbers: 43.30.Es, 43.35.Bf [SLB]

Pages: 1895–1910

I. INTRODUCTION

The acoustics of bubbly liquids is important to many areas of study, including acoustical oceanography, shallow water sonar, ship wakes, air/sea interfacial mass transfer, medical ultrasound, and nuclear reactor technology.^{1,2} Sound speed dispersion and attenuation in clean suspensions of bubbles is well understood when the acoustic excitation frequency is well below the resonance frequency of any bubble and the volume fraction of air is below a few percent. The volume fraction (VF) of air, or void fraction β , is defined as

$$\beta = V_g / V_m, \quad (1)$$

where V_g is the volume of the gas phase and V_m is the total volume of the mixture.

As the excitation frequency approaches the range of bubble resonance frequencies, the dispersion and attenuation grows rapidly. This effect is amplified further for bubble size distributions with sharp peaks. Commander and Prosperetti³ compared their propagation model to the experimental data available at the time and found poor agreement in such cases, even for void fractions as low as 3.7×10^{-4} . This has been attributed to the effects of multiple scattering between neighboring bubbles³ and experimental difficulties in measuring acoustic properties in the presence of high attenuation.⁴ Recently, a number of authors^{5–8} have put forward models that address multiple scattering. Unfortunately, experimental un-

certainty in the data that exists, and the absence of data for much of the parameter space, prevents an adequate evaluation of competing models.^{3,9} The fact that the data obtained by Silberman in 1957 is the most commonly cited data further emphasizes this situation.

There are two reasons for the scarcity of reliable experimental data for cases where resonance is important. (1) The attenuation, which is of $O(10)$ dB/cm and higher, prevents the use of traditional standing wave techniques. Even pulse propagation techniques fail because sound transmission beyond the near field of the source is not readily achieved. (2) Propagation in bubbly liquids is very sensitive to the void fraction, and to the bubble size and spatial distributions. A sufficiently accurate experimental determination and control of these parameters is difficult.

An approach designed to overcome these difficulties is described in the present work. In order to deal with the problem of high attenuation, measurement directly within the highly attenuating region of bubbly liquid was avoided. Instead, the frequency-dependent, complex impedance z_s at the surface of a layer of bubbly liquid was measured using a modified version of the standard two-sensor impedance tube.¹⁰ This was accomplished by measuring the acoustic pressures at two known positions within the bubble-free section of the impedance tube. Equations of plane wave acoustics^{10,11} were used to relate these acoustic pressures to the impedance at the interface between the bubble-free liquid and the bubbly liquid. This impedance is equivalent to the acoustic pressure/velocity ratio that would exist just inside the bubbly liquid if it were a homogeneous fluid. In the

^{a)}Current affiliation: University of Texas at Austin, Department of Mechanical Engineering; electronic mail: pswilson@mail.utexas.edu

present case, this measured impedance is equivalent to the effective impedance of the bubbly liquid and the effective propagation parameters are inferred from the measured impedance z_s . For the level of attenuation encountered in the present work, only plane progressive waves cross the measurement surface; therefore z_s is equivalent to the specific acoustic impedance $z = \rho c$. Since z is measured and the effective mixture density ρ is known, the phase speed and attenuation can be calculated from the complex sound speed c . Measurements of the phase speed and attenuation obtained in this way for a bubbly liquid composed of air and distilled water are presented. Void fractions ranging from 6.2×10^{-5} to 5.4×10^{-4} and bubble sizes centered around 0.62 mm in radius were investigated.

These measurements were compared with the model put forth by Commander and Prosperetti.³ The sensitivity of the model to input parameters is discussed in Sec. II. A description of the experiment is given in Sec. III. The results of a time-averaged experiment are discussed in Sec. IV. Finally, the results of a time-resolved experiment, used to investigate the effects of bubble population variability, are discussed in Sec. V.

II. OVERVIEW OF PROPAGATION MODEL

In 1989, Commander and Prosperetti³ brought together previously published results and composed a model for sound propagation in bubbly liquid which accounted for energy dissipation in a detailed way. Their final result is shown below for reference. Consider a host liquid with sound speed c_l , density ρ_l , viscosity μ , surface tension σ , and equilibrium pressure P_∞ . This liquid contains bubbles composed of a gas with thermal diffusivity D_g and ratio of specific heats γ . The bubbly liquid is treated like a homogeneous mixture in which the complex mixture sound speed c_m is given by

$$\frac{c_l^2}{c_m^2} = 1 + 4\pi c_l^2 \int_0^\infty \frac{a\varphi(a)da}{\omega_0^2 - \omega^2 + 2ib\omega}, \quad (2)$$

where ω is the circular excitation frequency and $\varphi(a)da$ is the number of bubbles per unit volume with equilibrium radius between a and $a+da$. The damping coefficient for a bubble of radius a driven at frequency ω is given by

$$b = \frac{2\mu}{\rho_l a^2} + \frac{P_{b,e}}{2\rho_l a^2 \omega} \text{Im} \Phi + \frac{\omega^2 a}{2c_l}, \quad (3)$$

where the three terms are due to viscous, thermal, and acoustic dissipation effects, respectively and $\text{Im} \Phi$ is the imaginary part of

$$\Phi = \frac{3\gamma}{1 - 3(\gamma - 1)iX[(i/X)^{1/2} \coth(i/X)^{1/2} - 1]}. \quad (4)$$

The bubble resonance frequency is

$$\omega_0^2 = \frac{P_{b,e}}{\rho_l a^2} \left(\text{Re} \Phi - \frac{2\sigma}{aP_{b,e}} \right), \quad (5)$$

where $\text{Re} \Phi$ is the real part of Φ . The definition,

$$X = D_g / \omega a^2, \quad (6)$$

TABLE I. Values of the physical parameters used for evaluation of Eq. (2).

$c_l = 1481 \text{ m/s}$	$D_g = 2.08 \times 10^{-5} \text{ m}^2/\text{s}$
$\rho_l = 998 \text{ kg/m}^3$	$P_\infty = 101.3 \text{ kPa}$
$\sigma = 0.0725 \text{ N/m}$	$\mu = 0.001 \text{ Pa s}$
$\gamma = 1.4$	

has been used and the quantity $P_{b,e}$ represents the equilibrium pressure in the bubble,

$$P_{b,e} = P_\infty + \frac{2\sigma}{a}, \quad (7)$$

where P_∞ is the hydrostatic pressure in the liquid. The wave number for propagation within the bubbly mixture can then be written from Eq. (2) as

$$k_m^2 = \frac{\omega^2}{c_l^2} + 4\pi\omega^2 \int_0^\infty \frac{a\varphi(a)da}{\omega_0^2 - \omega^2 + 2ib\omega}. \quad (8)$$

A. Sensitivity to small changes in bubble size

Small changes in bubble population parameters (or small uncertainties in their experimental determination) can have large effects on the (predicted) phase speed and attenuation in bubbly fluids. This will be demonstrated for different bubble size distributions using Eq. (2), assuming air bubbles in pure water at 20 °C and the corresponding physical parameters given in Table I.

Consider a monodisperse bubble population, given by

$$\varphi(a) = n\delta(a - \bar{a}), \quad (9)$$

where n is the number of bubbles per unit volume with equilibrium radius \bar{a} and the void fraction is

$$\beta = \frac{4}{3}\pi\bar{a}^3 n. \quad (10)$$

Substitution into Eq. (2) yields

$$\frac{1}{c_m^2} = \frac{1}{c_l^2} + \frac{4\pi\bar{a}n}{\omega_0^2 - \omega^2 + 2ib\omega}. \quad (11)$$

Phase speed V and attenuation coefficient A were calculated using Eq. (11) for void fraction $\beta = 10^{-4}$ and three bubble radii, $\bar{a} = 0.558, 0.62,$ and 0.682 mm. Both the phase speed and the attenuation curves, shown in Fig. 1, shift along the frequency axis as bubble size is varied. A small increase in peak attenuation is seen as bubble size decreases and the minimum phase speed increases slightly. The latter effect is not prominent on the logarithmically scaled axis, but is indeed present. The behavior of the attenuation for these parameters is governed by thermal dissipation, which is inversely proportional to the bubble radius squared.

Note that the pronounced effects shown in Fig. 1 are induced by only a 10% variation in bubble radius, or specifically, $\pm 62 \mu\text{m}$. The experimental ability to discern the difference between acoustic signals of 5.3 and 5.8 kHz (the frequencies of peak attenuation for the two smaller bubble sizes shown in Fig. 1) is often greater than our ability to discern the difference in the two corresponding bubble populations. If one were only interested in the size of a few

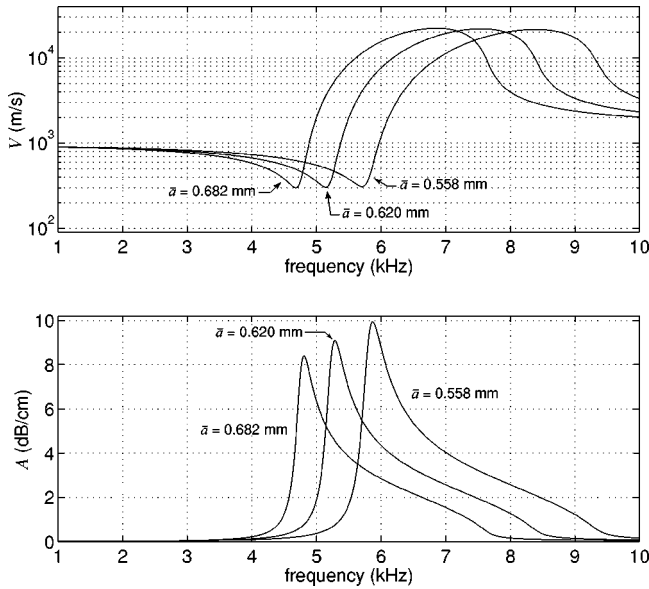


FIG. 1. The effect of the bubble size is demonstrated. Phase speed (upper) and attenuation (lower) is shown for void fraction $\beta = 10^{-4}$ and three bubble radii, which represent a $\pm 10\%$ change in radius.

bubbles this would not be the case; high magnification optical methods would provide sufficient resolution. However, when one endeavors to measure the sizes of hundreds of bubbles in motion, the statement bears a much greater significance.

B. Sensitivity to small changes in bubble size distribution width

The behavior of continuous bubble size distributions will now be considered. Foreshadowing what will be discussed in the experimental part of this work, the normal distribution will be explored. The probability density function of a normally distributed random variable is defined by

$$\varphi(a) = \frac{C}{s\sqrt{2\pi}} \exp\left[-\frac{(a-a_0)^2}{2s^2}\right], \quad (12)$$

where a_0 and s represent the mean value and the standard deviation of the distribution, respectively. Here, a_0 is taken to be the mean bubble radius, and C is used as a scaling parameter such that the void fraction β is given by

$$\beta = \frac{4\pi}{3} \int_{a_{\min}}^{a_{\max}} \varphi(a) a^3 da. \quad (13)$$

The purpose is to isolate the effect of increasing the standard deviation of the distribution, that is, the effect of widening the distribution.

Three distribution widths are considered, $s = 0.02, 0.03,$ and 0.04 mm, along with the following parameters: $\beta = 10^{-4}$, and $a_{\min} = 0.50, a_0 = 0.62, a_{\max} = 0.75$ mm. With $\varphi(a)$ thus described, Eq. (2) was evaluated numerically using an adaptive Simpson quadrature technique. The results are shown in Fig. 2. The peak attenuation varies about 1.8 dB/cm and the minimum phase speed varies by about 100 m/s. Again, the ability to acoustically detect the effects of

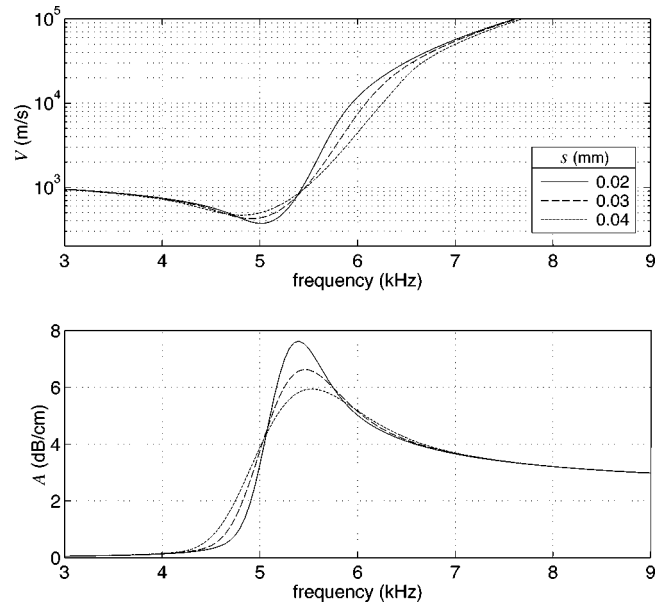


FIG. 2. The sensitivity to the distribution width is demonstrated. Phase speed (upper) and attenuation (lower) is shown for void fraction $\beta = 10^{-4}$ and three standard deviations s . The legend applies to both plots. Additional bubble population details are given in the text.

these distribution changes may be greater than the ability to detect and control the bubble population parameters in an experiment.

III. DESCRIPTION OF EXPERIMENT

An overview of the measurement procedure is given first and then the various subsystems will be described in greater detail. The bubbles were generated directly at the measurement plane of the impedance tube, by either a single needle, which was lowered down into the impedance tube, or with a bubble injection manifold (BIM) that fit inside the impedance tube and deployed multiple needles for higher void fractions. The needle and the BIM can be moved between the impedance tube and two other systems that were used to measure the bubble size distribution and the overall void fraction. The experimental procedure was composed of four subprocedures. The first subprocedure was the calibration of the impedance tube. The needle or BIM was then installed in an optical apparatus for the measurement of the bubble size distribution. When the needle was used, the void fraction was also measured in the optical apparatus. When the BIM was used, it was moved to a third apparatus for void fraction measurement. Finally, the needle or BIM was installed in the impedance tube and the surface impedance of the bubble layer was measured.

A schematic of the impedance tube system is shown in Fig. 3. Its construction, calibration, and operation have been described previously.^{10,12} Briefly, the system consists of a heavy-walled stainless steel tube (5.178 cm inner diameter, 2.541 cm wall thickness) with two custom fabricated wall-mounted hydrophones. The system was designed to admit waves that are plane (to a high degree of approximation), despite some coupling between the fill material and the tube wall. Acoustic excitation was provided by a Kildare Corporation TP-400/A piston source at the lower end. The bubbly

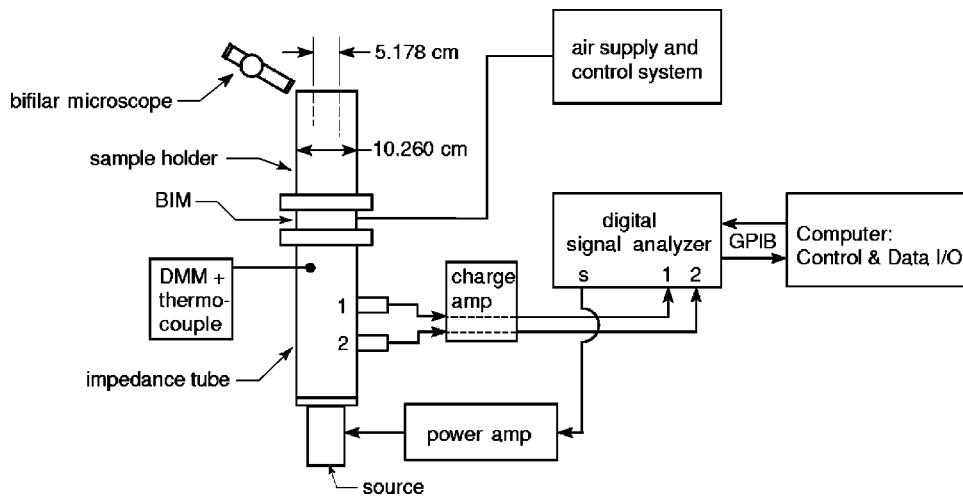


FIG. 3. The impedance tube system is shown in a schematic. The bifilar microscope is used during a calibration to measure the water height within the sample holder. The temperature of the impedance tube is monitored with a thermocouple and a digital multimeter. The remaining elements are discussed in the text.

liquid was contained within a sample holder mounted at the upper end. Signals were generated and received with a Hewlett-Packard HP89410A vector signal analyzer. The excitation signal was amplified by a Crown CE-1000 power amplifier and the hydrophone signals were conditioned with a Brüel & Kjær model 2692 charge amplifier. The transfer function between the two hydrophones was measured with the HP89410A and the data was transferred to a laptop computer for storage and processing. The effective plane wave surface impedance of the bubbly liquid is related to the measured transfer function using Eq. (1) of Ref. 10 and a set of calibration functions obtained in the calibration procedure.

A. Inversion procedure

In the present work, the high attenuation provided by the bubbly liquid suppressed reflections from the air–water interface at the top of the sample holder, giving it an effectively infinite length. Therefore, plane progressive waves crossed into the bubbly liquid and the complex sound speed (phase speed and attenuation) was available directly from the measured surface impedance. This condition existed for void fractions as low as $O(10^{-5})$, where on-resonance attenuation was $O(2)$ dB/cm. The sample holder used in these experiments was 14 cm in length and, therefore, minimum round trip attenuation was about 56 dB, which was sufficient to render the reflected signal indistinguishable from the background noise.

The phase speed and attenuation were obtained in the following way: Since $z_s = \rho c$ for the plane progressive waves incident on the surface of the bubbly liquid and $k = \omega/c$, the wave number k in the bubbly liquid is given by

$$k = \frac{\omega \rho}{z_s}, \quad (14)$$

where z_s is the measured impedance and ρ is obtained from the mixture density relationship $\rho = (1 - \beta)\rho_l + \beta\rho_g$. Finally, the phase speed is given by

$$V = \frac{\omega}{\text{Re}[k]}, \quad (15)$$

and the attenuation coefficient is given by

$$A = 20 \log_{10}(e) \text{Im}[k], \quad (16)$$

in dB/unit length.

The key assumption used in this analysis is that only a single plane wave mode exists in the sample holder at a given frequency. For waves propagating within a liquid-filled waveguide with elastic walls, at least two modes can be present at all frequencies. Both modes can have longitudinal and radial components, but the lowest-order mode approaches a true plane wave mode as the wall impedance increases.¹³ This impedance tube was designed such that when filled with bubble-free water, both the radial component of the plane wave mode and the higher-order mode are negligible and can be ignored.¹⁰ Below resonance, the bubbly liquid medium is acoustically soft in comparison to pure water and the elastic waveguide effects are minimized even further. Above resonance, the bubbly liquid is acoustically hard in comparison to pure water and the elastic waveguide effects are more prominent. An inversion scheme that properly accounts for both components of both modes does not exist and for now these elastic waveguide effects have been ignored. This will be discussed further in Sec. IV C.

B. Bubble production

The bubbly liquid consisted of air bubbles injected into distilled water using 30 gauge needles. The air was supplied from a compressed air bottle, through a primary regulator, then into a needle valve, a flow meter, and finally into the needle/s, either directly or through a manifold. A pressure gauge was connected between the flow meter and the needles. The flow rate and pressure measurements were used solely to ensure that the flow conditions, and hence the bubble population, remained as constant as possible throughout any given experiment. The bubble size distribution and overall void fraction were measured directly for each case.

For the first measurement presented in Sec. IV B, a single long needle (30 gauge, 0.31 mm diameter, 20.3 cm in length) was used. The needle was inserted into the sample holder from the open end of the tube and was positioned such that the bubbles were released at the measurement plane. Since changing the flow rate of air through a single needle changes the bubble size as well as the void fraction, one must add additional needles with the same per needle flow rate in order to independently increase the void fraction.

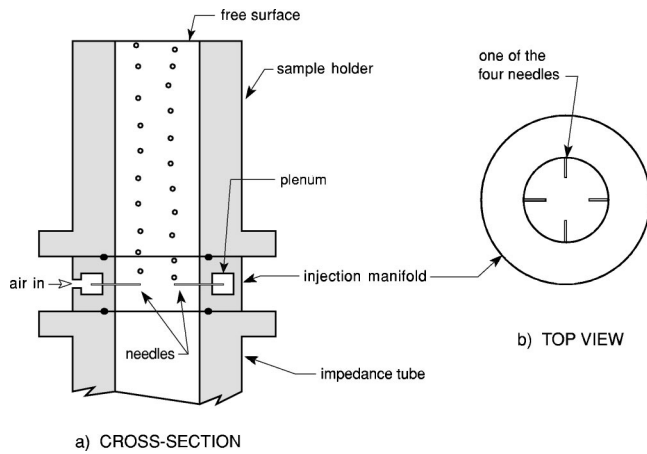


FIG. 4. A schematic diagram of the bubble injection manifold is shown. Part (a) is a cross-sectional view showing the injection manifold mounted between the impedance tube and the sample holder. Two needles are shown making bubbles. Part (b) shows a top view of the injection manifold separated from the impedance tube and the sample holder, with four needles in place. The needles were 30 gauge, 2.86 cm in length, and projected 1.27 cm into the tube.

Therefore, a bubble injection manifold (BIM) was developed to accurately position four 30 gauge needles at the measurement plane. The manifold body is of the same radial dimensions and material as the impedance tube itself, but contains an internal plenum that directs the air supply into the needles, as shown in Fig. 4(a). All the needles are coplanar with the impedance measurement plane. As the bubbles rose, they typically deviated from a directly upward path, yet they remained uniformly distributed along the longitudinal axis of the tube, and therefore the void fraction remained uniform in the direction of propagation. All but the first measurement presented in Sec. IV B were made using the BIM.

C. Measurement of bubble size

The bubble size distribution was determined by the direct photographic observation of the bubbles. Illumination was provided using a combination of direct and backlighting such that the outline of the bubbles appeared dark against a much lighter background. The single long needle or the BIM was placed within a small glass aquarium filled with distilled water. The height of water above the opening of the needles was the same as when installed in the impedance tube. The bubbles were free to rise up through the water column and they were photographed using a Kodak DC265 digital camera equipped with a macrolens and a strobe. The camera's maximum acquisition rate of about one photograph per 15 s was used. A machinist's scale located in the focal region was also photographed to provide a length reference. A schematic of the bubble size measurement system is shown in Fig. 5. Both the imaging and subsequent analysis were undertaken following guidelines published by The National Institute of Standards and Technology.¹⁴ A typical image obtained with this system is shown in Fig. 6(a) and the results of the image analysis are shown in Fig. 6(b). Note the relatively homogeneous distribution of bubbles both vertically and horizontally. A Cannon model GL1 digital video camera

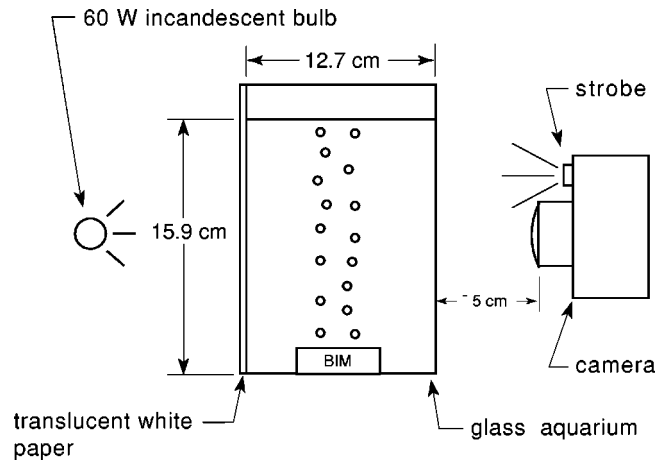


FIG. 5. A schematic diagram of the bubble size measurement system is shown with the BIM producing bubbles. The height of the water column is the same as that inside the sample holder during impedance measurements. Not shown is the machinist's scale used for a length reference.

was used to obtain photographic data at the higher rate required for the time-resolved measurements in Sec. V.

The bubbles exhibited a shape well approximated by an oblate spheroid. The images were processed with the image analysis software package NIH image. The major and minor axes $2b$ and $2c$ appeared in the plane of the image, and were determined by the NIH image in number of pixels. The length reference l_{ref} (length/pixel) was taken from the machinist's scale. The spheroidal volume is given by $V = \frac{4}{3} \pi b^2 c l_{\text{ref}}^3$. The effective spherical radius a was then calculated based on an effective spherical volume,

$$a = \left(\frac{3V}{4\pi} \right)^{1/3} = l_{\text{ref}} (b^2 c)^{1/3}. \quad (17)$$

The number of individual bubbles analyzed varied from experiment to experiment, but a typical number was 200.

Two primary factors that contribute to the uncertainty of the bubble size measurement are a deviation of the actual bubble shape from the model shape, and the limited length

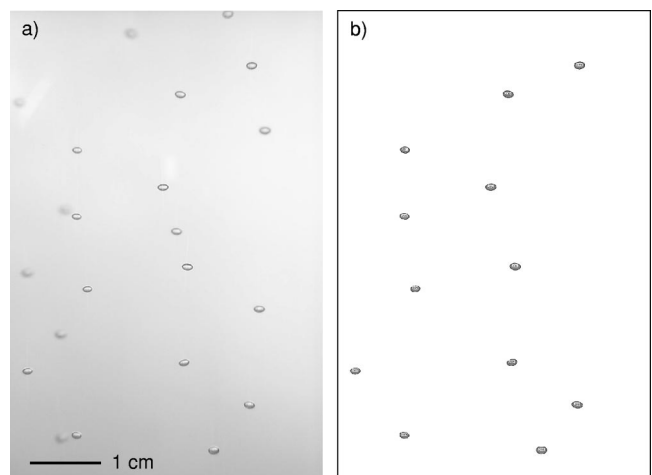


FIG. 6. A typical image obtained with the bubble size measurement system and four needles is shown in (a) along with a 1 cm scale bar. The same image after processing is shown in (b). The background was removed, each bubble was outlined, assigned a number, and the size data was stored in a file.

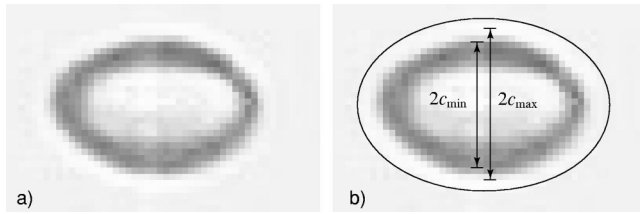


FIG. 7. An image of a typical bubble obtained by the imaging system is shown greatly enlarged in (a). Uncertainty due to pixelation is shown in (b). Equation (17) yields an effective spherical radius of 0.64 ± 0.05 mm for this bubble.

resolution of the pixelated images. Close scrutiny of a magnified image of a typical bubble reveals that the error due to pixelation is likely the greater of the two. This is demonstrated in Fig. 7. An image of a typical bubble used in this work is shown greatly enlarged in Fig. 7(a), so that individual pixels can be seen. The same bubble is again shown in part Fig. 7(b), but a slightly larger ellipse has been overlain and it is apparent that the shape is a good match to that of the bubble. There is some uncertainty about the boundary of the bubble, though. When the image was digitized, the pixels were assigned discrete values of brightness; therefore one cannot be sure exactly where the boundary was. The motion of the bubble during the exposure also contributes to this uncertainty. Consider the minor axis of the bubble, with length $2c$. If one assumes that the actual boundary lies somewhere between the outermost two pixels, then the minimum and maximum possible minor diameters are shown in the figure. For this bubble, $2c_{\min} = 20$ pixels and $2c_{\max} = 24$ pixels. Although not shown in the figure, a similar situation exists for the major axis, with $2b_{\min} = 32$ and $2b_{\max} = 36$ pixels, and for the length value, $1/l_{\text{ref}} = 230 \pm 2$ pixel/cm. Using Eq. (17) one finds the uncertainty in the measured effective spherical radius for this bubble is $a \pm 7.7\%$.

At this point a tradeoff is evident. For a given imaging system, one can increase the magnification and thereby increase the length resolution, but then the number of bubbles per image and the size of the statistical population is decreased. For fixed magnification, the uncertainty in bubble size is dependent on the bubble size. The uncertainty reported above was typical for this work. The maximum uncertainty in effective spherical bubble radius and did not exceed $a \pm 11\%$.

D. Determination of void fraction and low-frequency phase speed

For the single needle case, the bubbles were few enough in number that all could be counted individually and the total volume of air within the sample holder could be determined from the bubble size measurements.

When using the BIM, the void fraction was determined using a cylindrical acoustic resonator of length L , filled with the bubbly liquid of interest, and operating below the cutoff frequency of the first higher-order mode of the tube and well below the resonance frequency of the largest bubble. In this range, known as Wood's limit, the low-frequency mixture

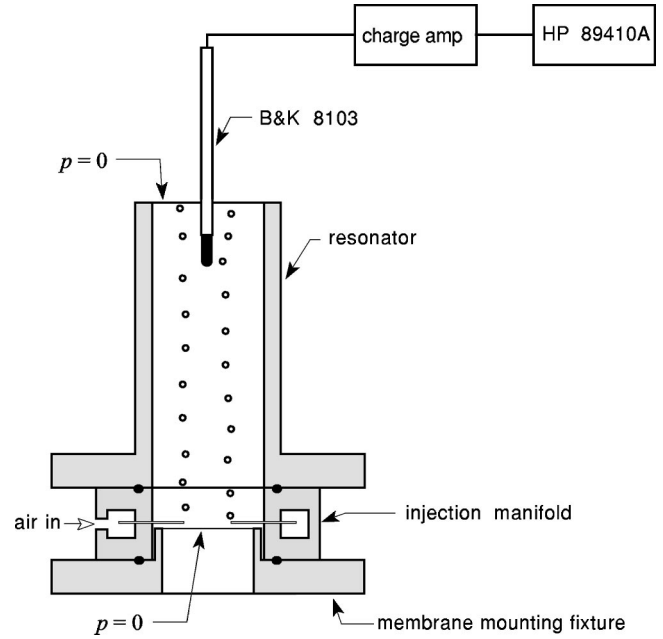


FIG. 8. A schematic diagram of the void fraction resonator is shown.

sound speed c_{mlf} is dependent on void fraction β , as given by the relation¹⁵

$$\frac{1}{c_{mlf}^2} = \frac{(1-\beta)^2}{c_l^2} + \frac{\beta^2}{c_a^2} + \beta(1-\beta) \frac{\rho_a^2 c_a^2 + \rho_l^2 c_l^2}{\rho_l \rho_a c_l^2 c_a^2}, \quad (18)$$

and c_a and ρ_a refer to the sound speed and density of air.

The resonator was modeled as a rigid-walled, open-open tube filled with a homogeneous liquid with sound speed c_{mlf} and supporting only one-dimensional plane-wave modes. The fundamental resonance frequency is given by $f_1 = c_{mlf}/2L$. The void fraction was determined from the measured frequency of the fundamental mode and Eq. (18). Sound speeds from the higher-order modes were also obtained. The modal sound speed was determined at each resonance frequency from $f_n = nc_n/2L$. Depending on the position of the hydrophone relative to the spatial acoustic field for each case, one, two, or three higher-order modes were observed.

A schematic diagram of the void fraction resonator is shown in Fig. 8. An aluminum tube (30.13 cm length, 5.08 cm diameter, 0.635 cm wall thickness) was mounted to the top of the air injection manifold and sealed with an o ring. The lower pressure release boundary condition was achieved by closing the tube with a thin mylar membrane that was glued to a mounting fixture. The mounting fixture interfaces with the air injection manifold, sealing the tube with an o ring, and positions the membrane just below the needles. A Brüel and Kjær type 8103 hydrophone was typically positioned 2 cm from the open end of the tube. The hydrophone signal was conditioned with a Brüel and Kjær type 2692 low noise charge amp, which also provided bandpass filtering between 10 Hz and 10 kHz. The natural resonances of the tube were excited by the noise created by the bubbles pinching off the needles, in a manner similar to that reported by

TABLE II. The table contains the results of the resonator-based void fraction measurements and the parameters used with Eq. (19) for calculation of the associated uncertainty. For Case 1, β was obtained from optical measurements and its uncertainty is discussed in Sec. IV B.

Case	1	2	3	4
$\beta \times 10^5$	5.49	33	41	54
$\Delta\beta$ (%)	...	± 3.4	± 3.2	± 3.3
L (cm)	...	31.5	31.5	31.5
ϵ_L (cm)	...	± 0.2	± 0.2	± 0.2
f (Hz)	...	857.5	810.0	730.0
ϵ_f (Hz)	...	± 6.875	± 6.125	± 6.000
ϵ_c (m/s)	...	± 7.8	± 7.1	± 6.7

Nicholas *et al.*¹⁶ The resulting hydrophone signal was acquired and acoustic spectra calculated with the HP89410A Vector Signal Analyzer.

Uncertainty ϵ_L in the length of the resonator and uncertainty in the resonance frequency ϵ_f due to the finite resolution bandwidth of the measured acoustic spectra lead to uncertainty ϵ_c in the measured sound speeds obtained from the resonator method and is given by

$$\pm \epsilon_c = \pm 2(L\epsilon_f + f\epsilon_L \pm \epsilon_L\epsilon_f). \quad (19)$$

In turn, this leads to an uncertainty range $\beta \pm \Delta\beta$ in the void fraction, obtained by the satisfaction of Eq. (18) for $c_{mlf} = 2Lf_1 \pm \epsilon_c$. Numerical values for these parameters are given in Table II.

E. Bubble population statistics

The bubble size data must be cast in the form of a probability density function $\wp(a)$ for input into the propagation model. Histograms of this data indicated that a truncated normal probability density function (PDF) was a suitable statistical model. Four parameters describe this distribution: the minimum radius a_{\min} , the maximum radius a_{\max} , the mean a_0 , and the standard deviation s . The bubble radii data, obtained as described in Sec. III C, were ranked in ascending order. The minimum and maximum radii values were used as a_{\min} and a_{\max} , respectively, and a_0 and s were estimated by fitting a model cumulative distribution function (CDF), to an empirical CDF obtained from the measured radii.

The measured or empirical CDF is given by

$$P_{\text{meas}}(a) = \frac{\bar{n}(a)}{N}, \quad a_{\min} \leq a \leq a_{\max}, \quad (20)$$

where $\bar{n}(a)$ is the number of observations with values less than or equal to a , and N is the total number of observations. The model CDF is given by

$$P(a) = \int_{a_{\min}}^a \wp(\xi) d\xi, \quad [0 \leq P \leq 1], \quad (21)$$

where \wp is PDF for the truncated normal distribution and is given by

$$\wp(a) = \frac{1}{s\sqrt{2\pi}} \exp\left[-\frac{(a-a_0)^2}{2s^2}\right], \quad a_{\min} \leq a \leq a_{\max}. \quad (22)$$

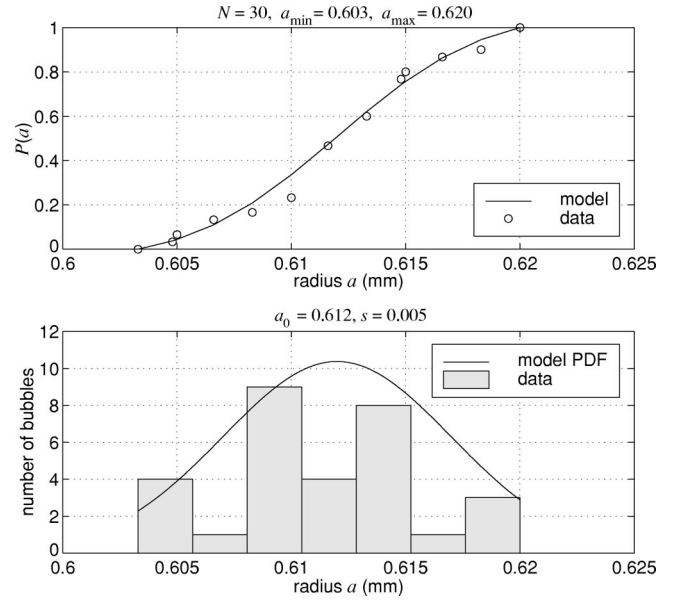


FIG. 9. The statistical data reduction is shown for bubbles created by a single needle. The upper frame shows the empirical CDF and the best-fit model CDF. The lower frame shows the resulting PDF and a histogram of the data, shown only for reference. The resulting distribution parameters are also shown (all radius units are in mm).

Model CDFs were then computed for a range of a_0 and s values, and a two-dimensional minimization of the chi-squared parameter was performed, yielding the best-fit mean and standard deviation. The chi-squared parameter χ^2 is given by¹⁷

$$\chi^2 = \sum_{i=1}^N \frac{(P_{\text{meas}}(a_i) - P(a_i))^2}{P(a_i)}, \quad (23)$$

where a_i is the i th member of the rank ordered vector of size measurements. An example of this statistical analysis is shown in Fig. 9. The empirical CDF fitting method was used because it relied entirely on the data. The PDF can be fitted to a histogram of the data, but this requires an arbitrary choice to be made for the histogram bin size.

F. Additional experimental details

Two different types of experiments were conducted, time-averaged, and time-resolved. In both of these experiments, the goal was to measure the frequency-dependent phase speed and attenuation in bubbly liquid in the resonance regime. In the original experimental design, it was assumed that the bubble population statistics and the void fraction would be stationary in time, as long as the observable air injection parameters remained constant. Subsequent experimental observations revealed that the bubble population statistics were not stationary, even when observable bubble production parameters were stable. This type of behavior exhibited by needle manifolds has been reported before.¹⁸ Further, moving the bubble injection manifold from the bubble sizing apparatus, to the void fraction resonator, and finally to the impedance tube, resulted in slightly different bubble distributions, even when the observable flow parameters did not change. Because the bubble population parameters were changing, a relatively small degree of fitting was

usually necessary in order to bring the time-averaged measurements and model into agreement, as discussed in Sec. IV. The discovery of nonstationarity lead to the idea that instantaneous, or time-resolved measurements should also be performed, as discussed in Sec. V.

IV. TIME-AVERAGED MEASUREMENTS

A. Procedure

The procedure for the experiment was composed of four parts: impedance tube calibration, bubble size distribution measurement, void fraction measurement, and finally impedance measurement.

For both the calibration and the impedance measurements, transfer functions between the signals from the two impedance tube hydrophones were measured using the HP89410A vector signal analyzer. Pseudorandom noise was used as the excitation signal over a frequency range of 5–9 kHz. The hydrophone signals were Hann windowed and processed to yield a frequency resolution bandwidth of 2.5 Hz. The calibration was performed first, using a procedure described in Ref. 10. Second, the single needle or the BIM was installed in the size distribution apparatus, shown in Fig. 5. The bubble size distribution was measured, as described in Sec. III C, which also yielded the void fraction for the single needle case. For the single needle case, bubble observation was limited to three photographs taken over a time frame of about 45 s, yielding 30 bubble images. When the BIM was used, 16 photographs were taken over a time frame of about four minutes yielding about 200 bubble measurements. The BIM was then installed in the void fraction apparatus shown in Fig. 8, and the void fraction was measured, as described in Sec. III D. This step also produced some additional phase speed data below 5 kHz.

Finally, either the single needle was inserted into the impedance tube, or the BIM was installed [as shown in Fig. 4(a)], and the tube was filled with distilled water. The bubble production was then initiated and the airflow rate and pressure were allowed to equilibrate. These flow parameters were kept constant throughout the size distribution, void fraction, and impedance measurement procedures. Because noise is produced by the bubble generation, the acoustic drive level was dependent upon the experiment. The coherence function between the two wall-mounted sensors was monitored with the vector signal analyzer. The minimum power amplifier gain necessary to obtain a near-unity coherence function was used. This ensured a sufficient signal-to-noise ratio and a linear bubble response.

The impedance measurement was then conducted by obtaining 100 spectral averages of the transfer function between the two wall-mounted sensors, using the signal analyzer. The averaged transfer function y was processed onboard the analyzer using

$$y = \frac{G_{12}(f)}{G_{11}(f)}, \quad (24)$$

where G_{12} is the averaged cross-spectrum given by

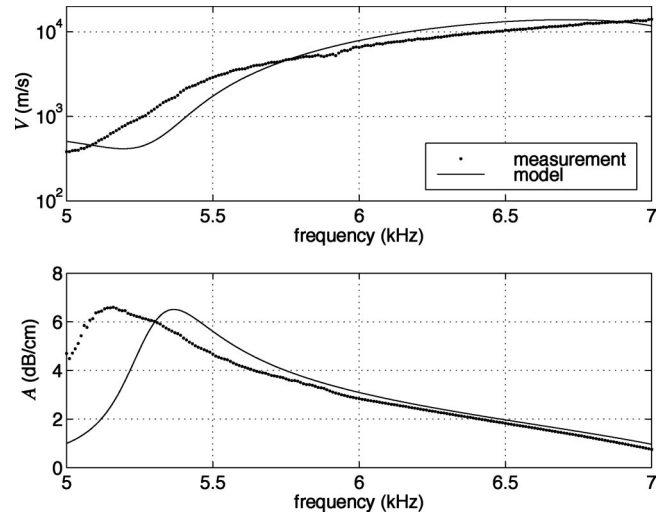


FIG. 10. Measured and predicted phase speed V and attenuation A for case 1, using the population parameters from Fig. 9.

$$G_{12}(f) = \frac{1}{N} \sum_{n=1}^N S_{2,n} S_{1,n}^*, \quad (25)$$

G_{11} is the averaged autospectrum, given by

$$G_{11}(f) = \frac{1}{N} \sum_{n=1}^N S_{1,n} S_{1,n}^*, \quad (26)$$

and individual spectra S_i are given by

$$S_i(f) = \text{FFT}[W(t)p_i(t)]. \quad (27)$$

In the above expressions, FFT is the fast-Fourier transform, $W(t)$ is the Hann window function, $p_i(t)$ is the acoustic pressure signal at the i th sensor position, the $*$ indicates the complex conjugate, and $N=100$ is the number of averages. Finally, the averaged transfer function y was sent to a computer over GPIB and the impedance was calculated off-line, as described in Ref. 10.

The above procedure was repeated for a total of four different void fractions, using a single long needle and four needles installed in the BIM. Note that when the single long needle was used, the BIM was not installed. The sample holder and the impedance tube were bolted directly together.

B. Results

The results for the single needle (case 1) are presented first. The results of the bubble population analysis is shown in Fig. 9. In the top part of the figure, the measured and the model CDFs are shown, along with the population size N , and the minimum and maximum bubble radii. In the bottom part of the figure, the best-fit model PDF is shown along with its parameters a_0 and s . A histogram of the bubble size data is also shown for reference, although it was not used in the analysis. The measured void fraction was $\beta = 5.49 \times 10^{-5}$.

The phase speed and attenuation measured for this bubbly liquid is shown in Fig. 10, along with the theoretical predictions given by Eq. (2), where $\varphi(a)$ is given by Eq. (12) and the four bubble population parameters from Fig. 9. A complete list of the parameters used in the production of Fig. 10 is given in Table III. The measured data and the

TABLE III. The table contains the parameters used to produce Figs. 10–15. All the physical parameters listed in Table I, except c_l , were also used. The bubble-free liquid sound speeds c_l were measured inside the impedance tube and show variation due to temperature. The presence of the asterisk (*) in a cell indicates that the parameter was fitted.

Case	1	1	2	3	4
Figure	10	11	13	14	15
$\beta(\times 10^{-5})$	5.49	6.20*	33	41	54
c_l (m/s)	1455.5	1455.5	1457.4	1455.8	1458.0
a_{\min} (mm)	0.603	0.627*	0.580*	0.580*	0.580*
a_0 (mm)	0.612	0.636*	0.600*	0.622*	0.638*
a_{\max} (mm)	0.620	0.645*	0.750*	0.710*	0.750*
s (mm)	0.005	0.005	0.031*	0.038*	0.035*

model are qualitatively similar, but they are shifted by about 2 kHz in frequency. Recall that the bubble size uncertainty is about $\pm 8\%$, and that the size data and the acoustic data are not taken at the same time. Taking this into account by adjusting the radius so that the frequency of the measured and predicted peak attenuation are the same yields Fig. 11, where the radii a_{\min} , a_0 , and a_{\max} were multiplied by 1.04. The volume-based void fraction estimate was also multiplied by $(1.04)^3$, yielding $\beta = 6.2 \times 10^{-5}$. Agreement is now much better across the entire frequency range. Because of the low attenuation at this void fraction, a reduced frequency range is shown. Outside of this range, the nonreflecting condition at the top of the tube was surely violated, which causes the inversion process to fail. Note that the data presented in this chapter has an inherent uncertainty associated with the impedance measurement of $\pm 2.2\%$.¹² Also note that other, less quantifiable sources of error are discussed in Sec. IV C.

The measurements obtained with the BIM are shown next. Four 30-gauge needles, 2.86 cm in length, were installed in the BIM for cases 2, 3, and 4. Each represents a different void fraction, which was achieved by slightly varying the overall flow rate. The needles were spaced 90° apart around the circumference of the BIM, and projected 1.27 cm into the tube. A single bubble size distribution measurement

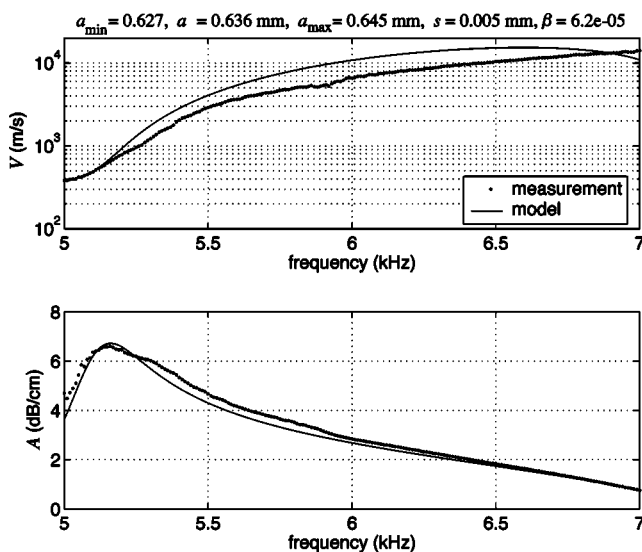


FIG. 11. The measured data for case 1 is shown again, but the predicted phase speed V and attenuation A was obtained by adjusting the bubble radius by +4%. The resulting bubble population parameters are also shown.

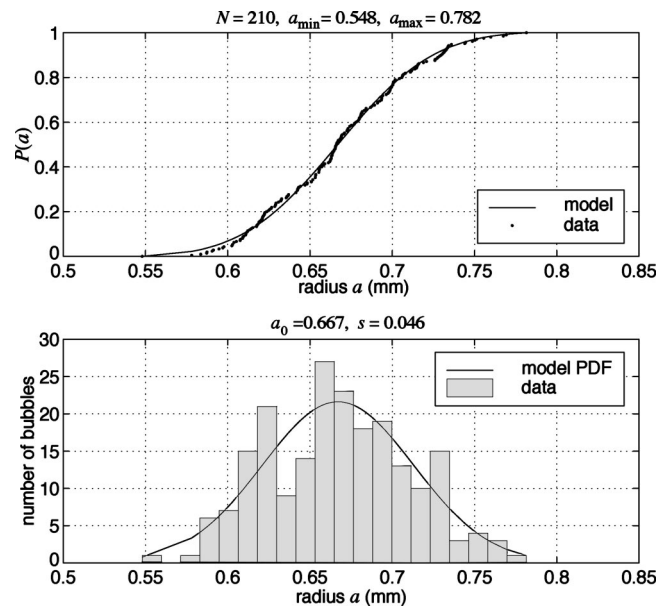


FIG. 12. The statistical data reduction is shown for bubbles created by the BIM with four needles and void fraction $\beta = 41.0 \times 10^{-5}$. The upper frame shows the empirical CDF and the best-fit model CDF. The lower frame shows the resulting PDF and a histogram of the data, shown only for reference. The resulting distribution parameters are also shown (all units in mm).

obtained with four needles and the intermediate flow rate is shown in Fig. 12. The measured phase speed and attenuation for cases 2, 3, and 4 are shown along with their associated predictions in Figs. 13, 14, and 15, respectively. A complete list of the model input parameters is given in Table III, and for convenience, the bubble population parameters are also shown in the figures.

In general, better agreement between the measurements and the model was found by fitting the bubble population statistical parameters within the uncertainty range of those measurements. Therefore, only the best-fit predictions are shown. Fitting the four bubble population statistics is easier

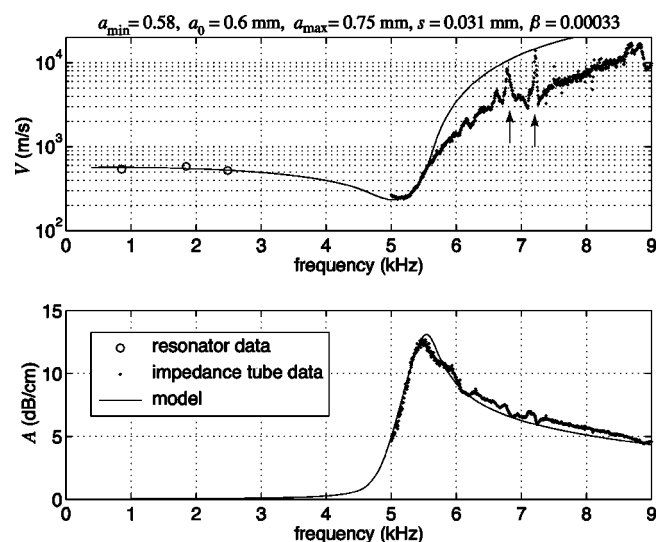


FIG. 13. The measured phase speed V and attenuation A for case 2 is shown. The predicted phase speed and attenuation that best fits the measured data is also shown along with the resulting bubble population parameters. The two vertical arrows near 7 kHz indicate sharp fluctuations in phase speed that are discussed in the text.

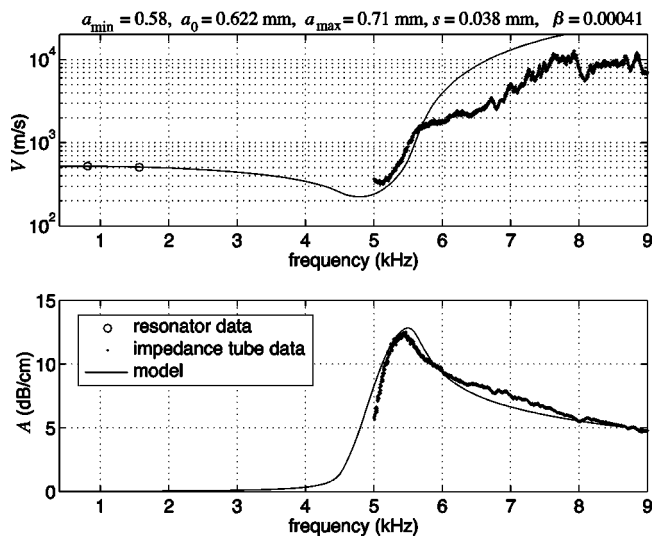


FIG. 14. The measured phase speed V and attenuation A for case 3 is shown. The predicted phase speed and attenuation that best fits the measured data is also shown along with the resulting bubble population parameters.

than one might expect because each parameter has a fairly unique effect. The fitting was done one parameter at a time, by hand. The difference between the data and the model in a certain range was minimized in a least squares sense by adjusting each parameter in turn. The mean bubble radius a_0 was fit using the frequency of peak attenuation. Standard deviation s was fit using the slope of the attenuation curve on the low-frequency side of the peak attenuation. Minimum bubble radius a_{\min} was fit using the width of the attenuation peak on the low-frequency side of the peak. Maximum bubble radius a_{\max} was fit using the width of the attenuation peak on the high-frequency side of the peak. No attempt was made to automatically fit all four parameters simultaneously to the absolute minimum error across the frequency range. The region around resonance was given priority.

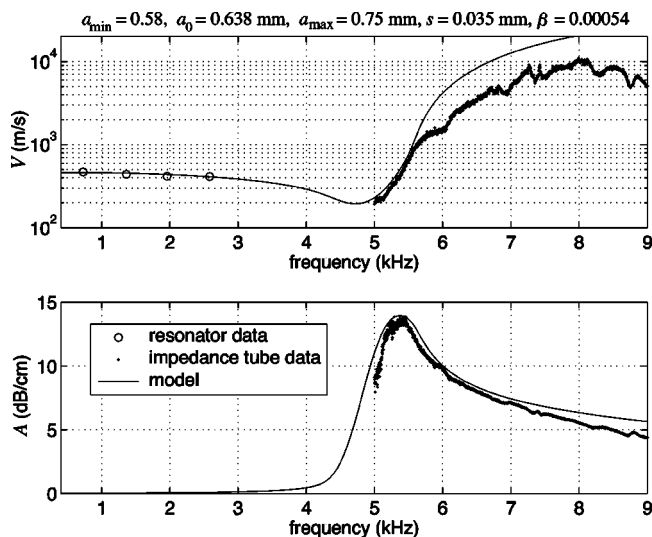


FIG. 15. The measured phase speed V and attenuation A for case 4 is shown. The predicted phase speed and attenuation that best fits the measured data is also shown along with the resulting bubble population parameters.

C. Discussion of time-averaged results

A number of general comments about the phase speed and attenuation data will be made and then these points will be discussed more fully. The predicted phase speeds above the resonance regime are generally higher than the measured values. The agreement is generally better for attenuation and is actually quite good. The deviation between the peak measured attenuation and the predicted value ranges from about 0.14 to 0.92 dB/cm. This agreement is better than that originally reported for peak attenuation in Ref. 3, where the minimum data/model deviation was about 2.5 dB/cm in Fig. 13 and was typically 10–15 dB/cm. The case 2–4 measurements display sharp fluctuations that case 1 does not. The fitted bubble population was always narrower than the measured bubble population. The low-frequency phase speed measurements obtained from the VF resonator (the open circles in Figs. 13–15) agreed well with the resonance regime phase speed measurements obtained from the impedance; the theoretical predictions tied them together.

1. Elastic waveguide effects

Elastic waveguide effects were mentioned in reference to the impedance data inversion procedure in Sec. III A. It has been shown that the phase speed in a liquid-filled waveguide with elastic walls is reduced relative to the liquid's intrinsic sound speed and is dispersive.¹³ As the frequency increases through resonance, the bubbly liquid changes from acoustically soft to acoustically hard (relative to the bubble-free water) and the elastic waveguide effects should become more pronounced. This could explain the mismatch between measurement and prediction above resonance. The effect was investigated theoretically by starting with an idealized case. Propagation in the impedance tube was modeled using an elastic waveguide dispersion relation from the literature [Eq. (5) in Ref. 13]. In this lossless model, the modal phase speeds of waves propagating in the internal liquid are determined as a function of frequency, waveguide geometry, and the intrinsic (free-field) sound speed of the internal liquid. A frequency of 7 kHz was used in the calculations. This frequency is above the bubble resonance frequency and a supersonic (relative to bubble-free water) phase speed is predicted for the unconfined bubbly liquid. This frequency is below the cutoff frequency for the first higher-order mode if this were a rigid-walled waveguide. For the elastic-walled case, two modes are predicted to propagate in the internal liquid and their phase speeds were calculated for a range of free-field sound speeds. Equation (5) of Ref. 13 is a lossless model, so the attenuation in the bubbly fluid was neglected. This is a major simplifying assumption, but the results are still useful for insight, if not absolute accuracy.

These results are presented in Fig. 16. The predicted phase speed of each of the two modes is plotted as a function of the intrinsic bubbly liquid sound speed. In other words, according to the elastic waveguide model, if the internal liquid sound speed was that given on the x axis; then the phase speed of each mode would be that given by the solid lines labeled “elastic mode 00” and “elastic mode 01.” The leading 0 indicates these are axisymmetric modes. If the waveguide was truly rigid, there would be only one mode and the

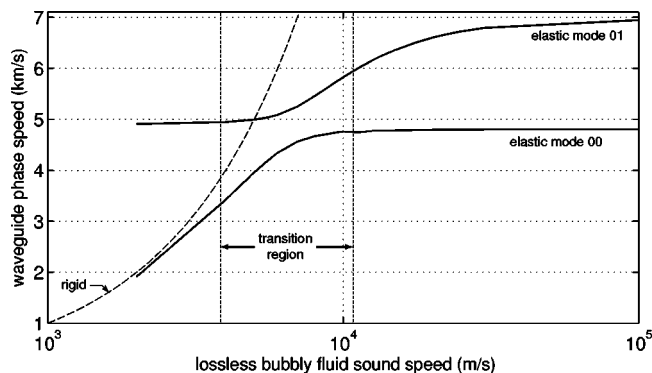


FIG. 16. The predicted phase speeds in an elastic waveguide are shown for supersonic regime bubbly fluid at 7 kHz. The mode labeled “00” is the plane wave mode, and “01” is the next higher-order mode. The dashed line represents the phase speed of the only mode present at this frequency for a rigid-walled waveguide.

waveguide phase speed would be the same as the intrinsic bubbly fluid sound speed. This case is represented by the dashed line labeled “rigid.” (The plot would be straight on a linear scale but is curved because the x axis is represented with a log scale.) The vertical dashed lines represent a transition region. Below this region, mode 00 has a phase speed close to the rigid waveguide sound speed and nearly plane wave fronts with little radial displacement. Mode 01 has curved wave fronts (concave down) and is not excited easily by plane disturbances. Above the transition region, both modes have curved wave fronts of almost the same concave up shape and similar radial components. Within the transition region, the phase speeds of the two modes are similar. Also notice that from midway through the transition region and higher in frequency, the two elastic mode speeds are always lower than the rigid-walled sound speed. This is a possible explanation why higher phase speeds are not measured, but it also brings up another issue.

According to the elastic waveguide model in Ref. 13, the 00 mode becomes nonplanar in the transition region. Now consider the situation within the impedance tube. An incoming plane wave begins to interact with the bubbly fluid at the interface between the tube and the sample holder and finds no suitable plane wave mode to excite. Part of the incident energy is reflected and the transmitted energy must get divided up into two propagating modes, and possibly a great number of evanescent modes. In such a case, the inversion that is used to extract phase speed and attenuation from the impedance measurements would break down, because it assumes only a single plane progressive mode for the transmitted wave.

Further evidence in support of this is available directly from the measured results. Consider the data in Fig. 11, which show a measured phase speed greater than 10 000 m/s at 7 kHz. According to Fig. 16, the maximum phase speed for mode 00 is less than 5000 m/s and about 7000 m/s for mode 01. This is still below cutoff for the next higher-order mode, so nothing should be able to propagate at 10 000 m/s. The existence of other modes must be interfering in such a way that an effective single mode phase speed is obtained from the inversion. An initial attempt was made using an

inversion technique that accounts for the higher-order mode, but it was not successful.¹² This remains an unresolved issue for phase speed measurements made with the present technique above resonance frequency, but plays a negligible role at resonance.

2. Signal and noise issues

The phase speed measurements exhibited an increase in sharp, rapid fluctuations with frequency when the BIM was used. The largest two of these features are indicated in Fig. 13 with small vertical arrows. The resulting increase in void fraction was accompanied by an increase in noise level produced as the bubbles broke off the needles. This noise represents energy entering the impedance tube from a direction not accounted for and, therefore, it could have corrupted the impedance measurements.

The measured coherence function was used to investigate this potential signal-to-noise issue. A coherence function greater than zero but less than unity implies one or a combination of the following:¹⁹ (1) The presence of extraneous noise in the measurements. (2) The output signal includes the result of excitation not measured by the input. (3) Some non-linearity between sensors 1 and 2. (4) Resolution bias errors are present.

Items (3) and (4) were ruled out because coherence functions obtained from measurements with a bubble-free transmission line termination were always close to unity, except in narrow bands. These bands are associated with nulls in the field that happen to be co-located with one of the pressure sensors. The acoustic drive level was always well below the threshold for nonlinear effects in distilled water; there were no bubbles between the measurement sensors. The resolution bandwidth of the measurement was not changed when bubbles were introduced.

Items (1) and (2) were possibilities, though, and an experiment was done to investigate this. Impedance measurements were conducted for five signal-to-noise ratios (SNR), in increments of 10 dB, with bubble production, and hence bubble-induced noise production constant. The SNR had no influence on the fluctuations. They did not increase as the SNR was decreased. At the lowest SNR, the coherence function was below 0.4 for much of the frequency range, but the rapid fluctuations did not deviate from the higher SNR cases. Therefore, excess bubble noise was ruled out as the cause.

What else could cause these perturbations? Perhaps it was the needles, or more specifically, the needles and the air inside. Case (1) shows none of these fluctuations. It was made with a single long needle lowered from the top. For the remaining cases, four needles were positioned parallel to and residing directly in the measurement plane. They present a larger cross section than did the long needle inserted from the top, and are positioned completely in the plane of interest. These are accompanied by an increased fluctuation level. This trend was observed to increase as the number of needles increased.¹² The presence of needles in the measurement plane certainly deviates from a homogenous bubbly liquid and could possibly cause the fluctuations.

These fluctuations could also be caused by smaller bubbles attached to the ends of the needles. Before a bubble

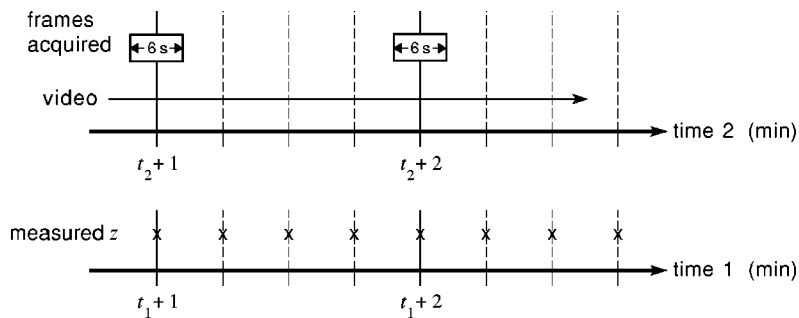


FIG. 17. Time lines for the time-resolved impedance measurements are shown. The impedance was measured in $66 \mu\text{s}$ segments once every 15 s. These measurement times are shown by the symbol “X.” The video recorded continuously, as indicated with the thin horizontal arrow. Individual frames were acquired from the video for six second blocks every minute.

detaches from its needle and rises upward, it spends some time attached to the end of the needle during the growth phase. During that time, it is certainly being interrogated by the incoming sound wave and therefore contributing to the impedance of the bubbly fluid. Because the bubble is interrogated multiple times, the net effect is to increase the influence that particular bubble size has on the final results. A multiple delta function bubble distribution present at the measurement plane could possibly account for the sharp phase speed peaks and fluctuations.

3. Bubble population parameters and void fraction

When the BIM was used (cases 2–4), the bubble radii standard deviation required to fit the propagation model to the impedance tube data was smaller than the standard deviation obtained from the photographic data. It was subsequently determined that the bubble population statistics varied on a time scale of about two minutes, even when the observable bubble production parameters were stationary. Therefore, bubble size distributions (BSDs) observed optically over several minutes were wider than the best-fit model BSDs used to match acoustic data gathered over about 45 s. Further, knowledge of the bubble population needs to be at least an order of magnitude more accurate to absolutely validate any propagation model. The bubble population parameter measurement should be conducted *in situ* with the impedance measurements. The low-frequency void fraction measurement technique proved to be reliable and agreement was found with both the resonance regime measurements and calculations. The *in situ* incorporation of the void fraction measurement would be ideal, but this task could be continued off-line. Finally, the increased control of bubble production would be beneficial, but if *in situ* population measurements are actually available, this is less important.

V. TIME-RESOLVED EXPERIMENT

An assumption made early in this work was that the experimental bubble populations would exhibit statistics that were stationary in time. As previously discussed, this was not the case. Since the bubble population characterization and the impedance measurement could not be done at the same time, an absolute comparison between the propagation model and experimental measurement was not possible. An investigation of the relative variability was possible and became the goal of the time-resolved experiment. In order to do this, the bubble population parameters and the acoustic impedance were sampled on similar time scales.

A. Procedure

The procedure for instantaneous impedance measurements was almost the same as that for the time-averaged measurements just described. Four subprocedures were again performed, but in the following order: calibration, void fraction measurement, impedance measurement, and then bubble size distribution measurement. The void fraction measurement procedures were unchanged from the time-averaged experiment.

The goal of this experiment was to observe the time variation of both the bubble size distribution and the propagation parameters. A periodic chirp was used as the excita-

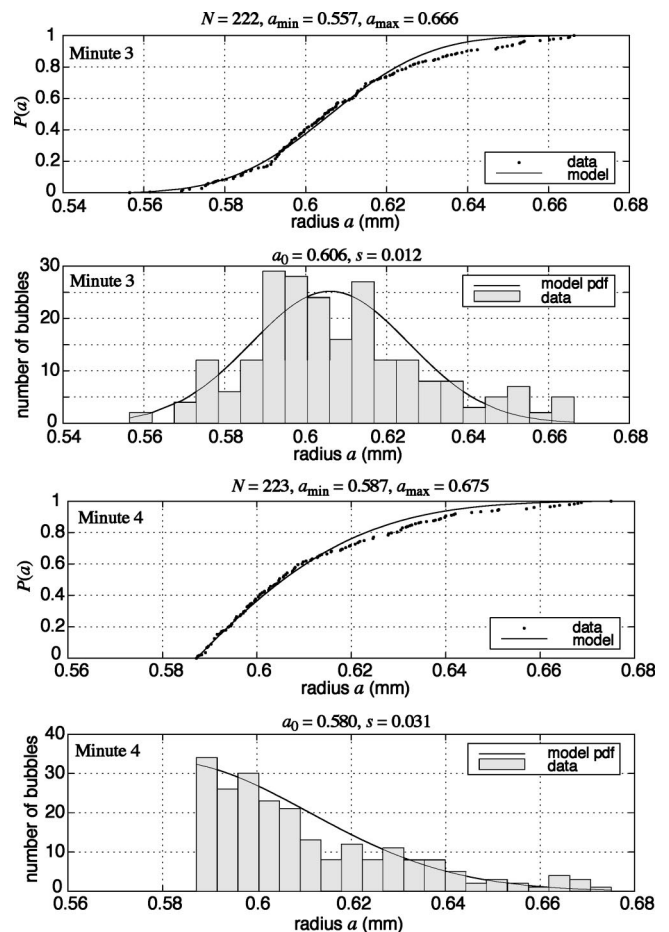


FIG. 18. Bubble population statistics at minutes 3 and 4 of the time-resolved experiment. For each case, the upper frame shows the empirical CDF and the best-fit model CDF. The lower frame shows the resulting PDF and a histogram of the data, shown only for reference. The resulting distribution parameters are also shown (all units in mm).

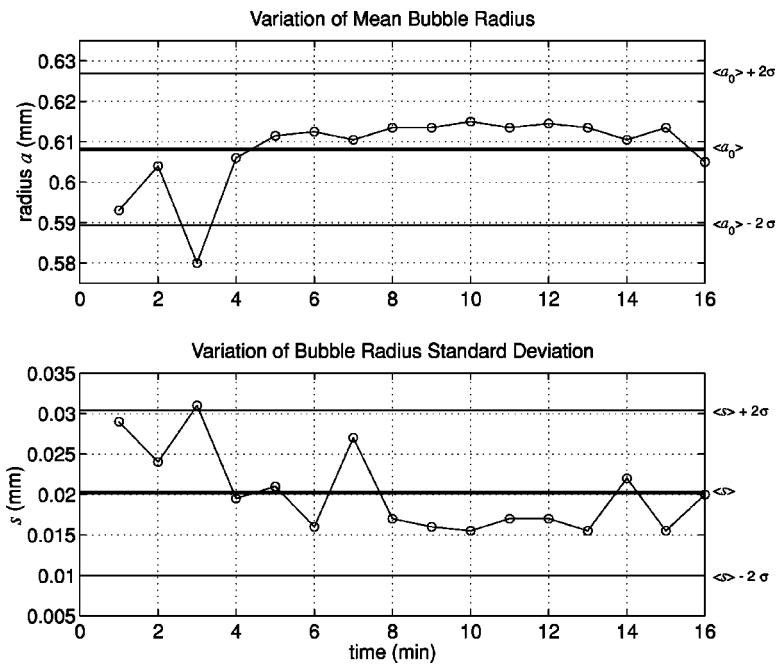


FIG. 19. Mean bubble radius a_0 and standard deviation s are shown as a function of time. The thick horizontal lines represent the global means, $\langle a_0 \rangle$ and $\langle s \rangle$, and the thin horizontal lines represent two standard deviations above and below the global means.

tion signal because it requires no averaging and could be used to obtain near-instantaneous measurements. The signal was periodic with respect to the data acquisition time window and was 66 ms in length. A Hann window was used, which resulted in a resolution bandwidth of 15 Hz for a frequency range of 4–10 kHz. The data from 4–5 kHz and 9–10 kHz was discarded.

The calibration was performed using the periodic chirp, then the void fraction measurements were performed. The impedance measurement procedure began as it did in the previous section, but instead of taking 100 back-to-back averages, one chirp-excited impedance measurement was taken every 15 s for 16 min. This was done for a single void fraction case, using the BIM and four needles.

The still camera could not operate fast enough to sample the bubble size distribution on the same time scale as the impedance, so a digital video camera was used. The resolution of this video camera was not high enough to image many bubbles in the same frame, so high magnification was used and one or two bubbles from a single needle were imaged instead. After the bubble production was initiated and the flow parameters were stabilized, video was recorded continuously for 16 min, but only 6 s blocks taken every 60 s were processed. Therefore, 6 s of video for each event translated into 180 frames and about 220 bubbles per event. Under such circumstances, the bubble population is sampled photographically for bubble size in 33 1/3 ms intervals, and the impedance measurements sample 66 ms of the bubble population. A time line relating the impedance measurements and the size measurements is shown in Fig. 17. These images were all processed for bubble size, as described in Sec. III C.

B. Results

Using the technique given in Sec. III E, the bubble population statistics were obtained for each of the 16 populations. The CDF/PDF plots for the entire dataset are presented in Ref. 12 along with additional details about the data acquisi-

tion and analysis. The uncertainty in these bubble size measurements was $\pm 11\%$. The histograms of two neighboring bubble populations, generated one minute apart, are presented in Fig. 18. It is clear that the bubble population is changing dramatically on this time scale. The statistical bubble population parameters are plotted as a function of time in Figs. 19 and 20. The following notation and definitions were used. The global mean of a variable x is given by $\langle x \rangle$, where

$$\langle x \rangle = \frac{1}{16} \sum_{i=1}^{16} x(t_i), \quad (28)$$

and the standard deviation of x is given by

$$\sigma = \left\{ \frac{1}{16-1} \sum_{i=1}^{16} [x(t_i) - \langle x \rangle]^2 \right\}^{1/2}. \quad (29)$$

The variation of the mean bubble radius $a_0(t)$ and the radius standard deviation $s(t)$ is shown in Fig. 19. It may appear that a start-up transient was captured at the beginning of the upper plot, but recall that the observable control parameters of the bubble production process had reached steady state prior to the image acquisition and were not fluctuating. The curves shown in Figs. 19 and 20 confirm that the bubble population parameters are fluctuating dramatically and rapidly. Variation greater than two standard deviations can occur within a minute, and there is also evidence of long term trends, specifically in the standard deviation.

The void fraction, determined just prior to the impedance measurements using the resonator technique, was 3.9×10^{-4} . Phase speed V and attenuation A were obtained from each impedance measurement. For compatibility with the bubble size data, V and A for four consecutive measurements were averaged together to represent one minute of time, then the next four were averaged, and so on, until 16 min of measured V and A were obtained. Predicted phase speeds and attenuations were then calculated using Eqs. (2)

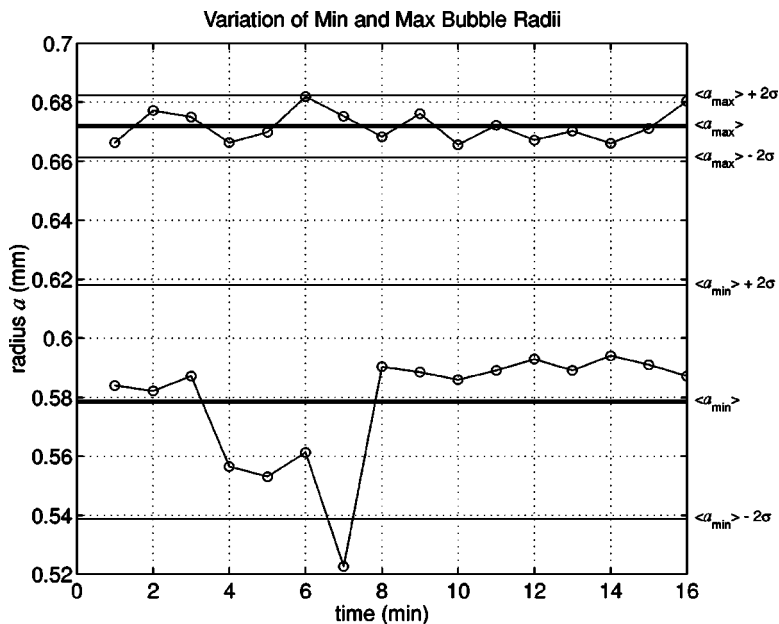


FIG. 20. Minimum and maximum bubble radius is shown as a function of time.

and (12), the measured bubble population parameters from Figs. 19 and 20, and the measured void fraction.

In order to compare the variability of these multifaceted measurements and calculations, attention will be focused on two parameters: the maximum attenuation $A_{max}(t)$ and the frequency at which that maximum attenuation occurs, $f_m(t)$. Since the relative variation is of interest, the results will be presented in normalized form, $A_{max}/\langle A_{max} \rangle$ and $f_m/\langle f_m \rangle$, where the brackets represent the global mean as defined in Eq. (28), and the normalized standard deviation of parameter x is given by

$$\bar{\sigma} = \left\{ \frac{1}{16-1} \sum_{i=1}^{16} \left[\frac{x(t_i)}{\langle x \rangle} - 1 \right]^2 \right\}^{1/2}. \quad (30)$$

The results are given in Figs. 21 and 22 for attenuation and

frequency, respectively. When cast in this form, the predicted variation of the acoustic properties due to changes in the bubble population is quite similar to the observed variation. For the normalized frequency parameter, a $\bar{\sigma}$ variation of 0.87% is predicted and 0.84% is observed. For the normalized attenuation parameter, a $\bar{\sigma}$ variation of 4.9% is predicted and 5.3% is observed.

C. Discussion

The similarity between predicted and observed statistical variability of the peak attenuation and its frequency is quite good. Making comparisons of the relative, as opposed to absolute quantities minimized the problems relating to the

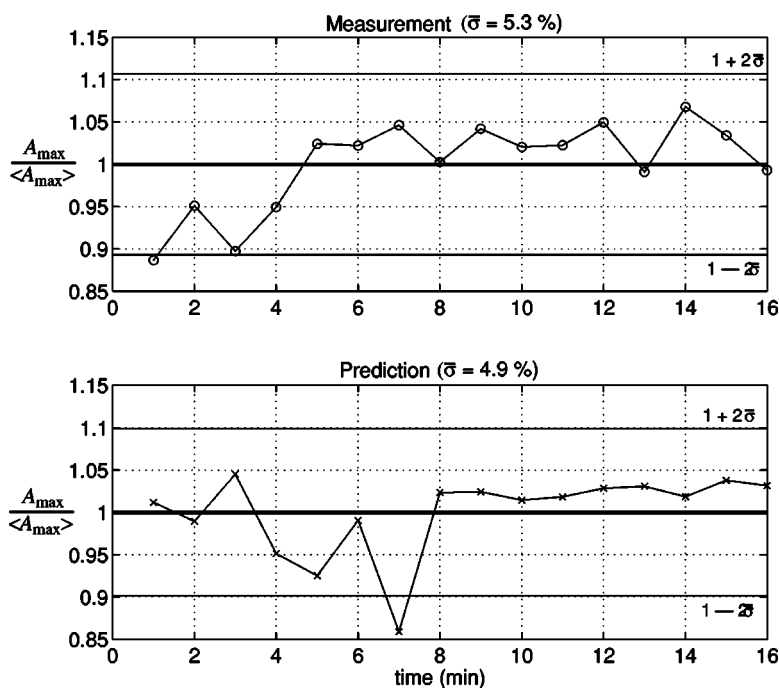


FIG. 21. Measured and calculated maximum attenuation is shown as a function of time, normalized by the global mean.

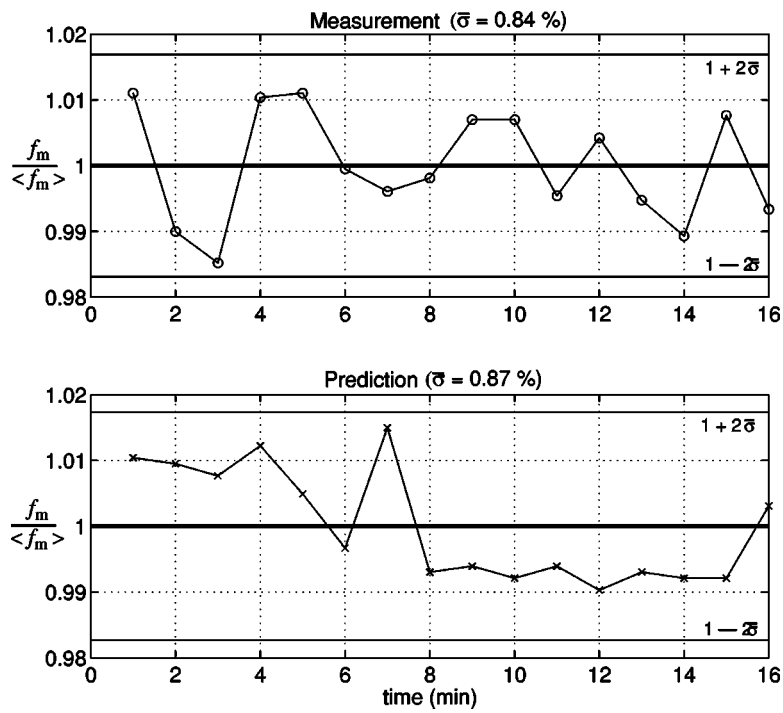


FIG. 22. Measured and calculated frequency of maximum attenuation is shown as a function of time, normalized by the global mean.

uncertainty of the instantaneous bubble population, including the uncertainty in the absolute size of the bubbles.

In the region around peak attenuation, the phase speed is near its minimum, hence the elastic waveguide effects are minimized. Therefore, the inversion problems that are evident well above resonance play no role here. In addition to the inherent errors already discussed, there is an additional source of bias error for this experiment. The resolution bias error was made negligible in the time-averaged experiment, as confirmed by near unity measured coherences. The short excitation time in the present experiment lead to a resolution bandwidth of 15 Hz, which increased the maximum estimated resolution bias errors a small amount. The level of error is still significantly smaller than the observed fluctuation, which can confidently be attributed to changes in the medium itself.

It has long been predicted from theory that propagation in bubbly fluids is sensitive to the bubble population parameters. The present work has reinforced this notion both theoretically and experimentally. The significance of these results is the following: Existing theory, specifically that represented by Eq. (2), is capable of predicting the *relative variability* of sound propagation parameters in bubbly liquids at resonance, if the bubble population parameters are known, up to a void fraction of 5.4×10^{-4} . This appears to be the case, despite the fact that the *absolute accuracy* of the model represented by Eq. (2) has not been verified experimentally.

VI. SUMMARY AND CONCLUSIONS

In this work, the acoustics of bubbly liquid was investigated. Experiments were conducted that focused on linear plane wave propagation parameters in the frequency range surrounding individual bubble resonance. An impedance tube system was used to measure the phase speed and attenuation in fluids composed of air bubbles and distilled water for a

range of void fractions between 6.2×10^{-5} and 5.4×10^{-4} . The bubble size distributions were centered around 0.62 μm in radius. The void fraction and size distribution parameters were obtained from secondary measurements done just prior to or just after the measurement of impedance. It was found that the bubble population statistics were not stationary in time. They were changing during the experiments, even though the observable bubble production controls were stationary in time. Despite this discrepancy, it was found that the model discussed in Sec. II could be used to describe measurements near resonance. In order to do so, some of the bubble population parameters were used to fit the model to the measurements. This fitting was within the uncertainty of the population parameter measurements and yielded good agreement between the model and the observations near resonance for both phase speed and attenuation. This agreement is better than the agreement obtained between the model and measurements previously reported.³ Above resonance, the model overpredicted the observed phase speed, however, it was determined that the inversion technique used to obtain the phase speed and attenuation from the impedance measurements was not appropriate for use when the phase speed rose above approximately 2000 m/s. Therefore, further use of the impedance tube technique for an investigation of the supersonic regime above resonance will require additional work on the inversion process.

In a second round of experiments, the *variability* of both the bubble population statistics and the propagation parameters were investigated. The population statistics measurements and impedance measurements were still done in succession, but, here, the time scales of the two sets of measurements were much more similar than in the time-averaged experiment. It was found that Eq. (2) could indeed predict the *relative variability* of the propagation parameters at resonance frequency, specifically the peak attenuation and

the frequency at which it occurred. This is the case despite the fact that the *absolute accuracy* of Eq. (2) has not been experimentally verified. Another general conclusion drawn from these results is the following: the observation of acoustical parameters in bubbly liquids must be done on the same time scale as the observation of the associated bubble population parameters.

Finally, these results indicate that the *absolute accuracy* of Eq. (2) and of other competing theories can be verified near resonance using the existing impedance tube system. In order to do this, a colocated instrument to simultaneously measure bubble population statistics, or more stable bubble production equipment is required. In either case, the resolution and accuracy of the sizing apparatus must be increased by an order of magnitude over the present system, and the number of bubbles observed per unit time should also be increased by an order of magnitude.

ACKNOWLEDGMENTS

This work was supported by the U.S. Navy Office of Naval Research Ocean Acoustics Program and Boston University. The authors acknowledge the efforts of Eun-Joo Park, who carried out the bubble size measurements and analysis presented in Sec. V.

¹T. G. Leighton, *The Acoustic Bubble* (Academic, London, 1994).

²H. Medwin and C. S. Clay, *Fundamentals of Acoustical Oceanography* (Academic, Boston, 1998).

³K. W. Commander and A. Prosperetti, "Linear pressure waves in bubbly liquids: Comparison between theory and experiments," *J. Acoust. Soc. Am.* **85**, 732–746 (1989).

⁴E. Silberman, "Sound velocity and attenuation in bubbly mixtures measured in standing wave tubes," *J. Acoust. Soc. Am.* **29**, 925–933 (1957).

⁵Z. Ye and L. Ding, "Acoustic dispersion and attenuation relations in bubbly mixture," *J. Acoust. Soc. Am.* **98**, 1629–1636 (1995).

⁶C. Feuillade, "The attenuation and dispersion of sound in water containing multiply interacting air bubbles," *J. Acoust. Soc. Am.* **99**, 3412–3430 (1996).

⁷F. S. Henyey, "Corrections to Foldy's effective medium theory for propagation in bubble clouds and other collections of very small scatters," *J. Acoust. Soc. Am.* **105**, 2149–2154 (1999).

⁸S. G. Kargl, "Effective medium approach to linear acoustics in bubbly liquids," *J. Acoust. Soc. Am.* **111**, 168–173 (2002).

⁹S. A. Cheyne, C. T. Stebbings, and R. A. Roy, "Phase velocity measurements in bubbly liquids using a fiber optic laser interferometer," *J. Acoust. Soc. Am.* **97**, 1621–1624 (1995).

¹⁰P. S. Wilson, R. A. Roy, and W. M. Carey, "An improved water-filled impedance tube," *J. Acoust. Soc. Am.* **113**, 3245–3252 (2003).

¹¹V. Gibiat and F. Laloë, "Acoustical impedance measurements by the two-microphone-three-calibration (TMTC) method," *J. Acoust. Soc. Am.* **88**, 2533–2545 (1990).

¹²P. S. Wilson, "Sound propagation in bubbly liquids," Ph.D. dissertation, Boston University, 2002.

¹³L. D. Lafleur and F. D. Shields, "Low-frequency propagation modes in a liquid-filled elastic tube waveguide," *J. Acoust. Soc. Am.* **97**, 1435–1445 (1995).

¹⁴A. Jillavenkatesa, S. J. Dapkunas, and L.-S. H. Lum, *Particle Size Characterization*, U.S. Government Printing Office, Washington DC, 2001.

¹⁵A. B. Wood, *A Textbook of Sound*, 2nd ed. (G. Bell and Sons, London, 1941).

¹⁶M. Nicholas, R. A. Roy, L. A. Crum, H. Oğuz, and A. Prosperetti, "Sound emissions by a laboratory bubble cloud," *J. Acoust. Soc. Am.* **95**, 3171–3182 (1994).

¹⁷J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures* (Wiley-Interscience, New York, 1971).

¹⁸J. A. Schindall, "Acoustic scattering from compact bubble clouds," Ph.D. dissertation, The University of Mississippi, 1995.

¹⁹J. S. Bendat and A. G. Piersol, *Engineering Applications of Correlation and Spectral Analysis*, 2nd ed. (Wiley, New York, 1993).

A phase-space Gaussian beam summation representation of rough surface scattering

Goren Gordon and Ehud Heyman

Department of Physical Electronics, Tel Aviv University, Tel Aviv 69978, Israel

Reuven Mazar^{a)}

Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

(Received 29 June 2004; revised 14 December 2004; accepted 23 January 2005)

A Gaussian beam (GB) summation representation for rough surface scattering is introduced. In this scheme, the coherent and incoherent scattered fields are described by a phase-space summation of GBs that emanate from the rough surface at discrete set of points and directions. It thus involves stochastic GB2GB scattering matrices for the coherent and incoherent fields, and deterministic GB propagators. It benefits from the simplicity and accuracy of the latter, and can be used in applications involving propagation in complex scenarios comprising inhomogeneous media with rough surface boundaries. The GB2GB matrices are calculated from the statistical moments of the scattering amplitude, which are given either analytically or empirically. An analytical and numerical example for weakly rough surface is presented and discussed. Applications to the more complicated propagation scenario of doubly rough surface waveguide with multiple reflection phenomena will be presented in a follow-up publication. © 2005 Acoustical Society of America.

[DOI: 10.1121/1.1869712]

PACS numbers: 43.30.Hw, 43.20.Fn [LLT]

Pages: 1911–1921

I. INTRODUCTION

Wave propagation in complex environments containing rough surface boundaries represents one of the main topics both of electromagnetics (Beckmann and Spizzichino, 1987) and acoustics (Ellis *et al.*, 1993; McDaniel, 1993) (see Etter, 2001, 2003 for references in underwater propagation). Rigorous analysis of rough surface scattering is based on plane-wave spectrum (Voronovich 1999), yet tracking of the plane-wave basis function through interaction with rough surfaces is rather complicated and computationally inefficient because, in addition to the specular rays, there are multiple nonspecular incoherent contributions originating from different spatial locations of the rough surfaces, thus leading to highly nonlocal integral representations. These difficulties become critical when the scattered field needs to be tracked through complex inhomogeneous medium, since the plane waves need to be adapted locally to the medium inhomogeneity. Localization of the rough surface scattering and of the propagation model can be affected asymptotically via ray theory: Prescriptions for how to adopt the ray theory for analyzing high-frequency phenomena in such structures are summarized in the stochastic geometrical theory of diffraction (SGTD), which has been formalized as an extension of its deterministic counterpart (GTD), when the parameters of the medium exhibit random spatial and temporal variations (Mazar and Felsen, 1989). The difficulties of the ray formulation stem from the nonuniform asymptotic structure of the ray field that fails near focal and caustic regions. Furthermore, a proper coverage of the propagation domain requires

the tracking of a dense lattice of rays, which becomes computationally intensive in complicated medium.

To circumvent these difficulties, in this work a new method based on the frame-based Gaussian beam summation (GBS) is proposed. In the GBS method, the source field is expanded into a spectrum of GBs that emanate from the source domain in all directions and thereafter are tracked locally in the medium. Unlike the plane-wave basis functions, GBs can be easily and efficiently propagated in inhomogeneous media and, unlike ray representations, they are insensitive to caustic and foci regions. Furthermore, when properly formulated, the GBS does not require a dense set of beams. Thus, the GBS representation combines the localization properties and algorithmic ease of the ray representation with the uniform features of the plane-wave representation.

There are essentially two classes of GBS representations, one that addresses point source configurations, e.g., Červený *et al.*, 1982; Babich and Popov, 1989; Porter and Bucker, 1987; Porter, 1997, and one that deals with distributed (or aperture) source configurations, e.g., Bastiaans, 1980; Steinberg *et al.*, 1991; Lugara *et al.*, 2003; Shlivinski *et al.*, 2004 [see Figs. 1(b) and (a), respectively]. In this paper we apply the latter to the problem of rough surface scattering. Originally, this scheme was based on Gabor's series expansion (Bastiaans, 1980; Steinberg *et al.*, 1991); hence, it suffered from certain inherent difficulties of this representation, but later on it was reformulated using the theory of windowed Fourier transform (WFT) frames (Lugara *et al.*, 2003; Shlivinski *et al.*, 2004), which remedied these difficulties (see an overview in Shlivinski *et al.*, 2004). The theory now provides a rigorous phase-space framework that decomposes any field in an aperture of a real or virtual source to a set of simple GBs that emanate from a discrete set of points

^{a)}Deceased.

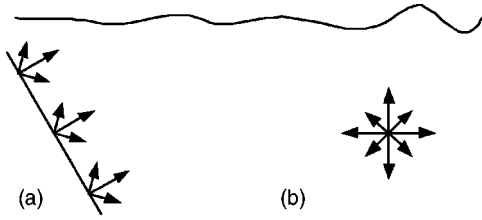


FIG. 1. Beam expansion of the source field field. (a) A phase-space beam expansion of an aperture source. (b) An expansion of a point source to GBs emerging from the source in a discrete set of directions.

in the aperture into a discrete set of directions.

The implementation of the GB procedure in the present rough surface scattering application is managed by a GB-to-GB (GB2GB) scattering matrix, where the incident field is decomposed into GBs, and each GB is then scattered into a discrete spectrum of GBs. This scattering matrix is the Gaussian equivalent of the scattering cross section of plane waves in the conventional spectral analysis of rough surface scattering. It can now be used in the overall problem of an inhomogeneous medium with rough surface boundaries, by propagating the GBs easily within the medium and using the GB2GB scattering matrix to acquire the relevant GBs scattered from the rough surface.

Concerning the presentation layout, we start with brief overviews of rough surface formulation and of the frame-based GBS (Secs. II and III A, respectively). In Sec. IV we combine the two to analyze rough surfaces using the GBS method, and provide general definitions for the coherent and the noncoherent GB2GB scattering matrices. An analytical example of a weakly rough surface, for which the scattering matrices can be calculated in closed form via the small perturbation approximation, is presented in Sec. V, together with a numerical example that demonstrates the phase-space footprints of the rough surface scattering processes. The presentation ends in Sec. VI with concluding remarks.

II. ROUGH SURFACE FORMULATION

A. Spectral representation

We consider the wave field in the lower half-space $z > h(\mathbf{x})$, with an upper rough surface $z = h(\mathbf{x})$. Here, z is the depth coordinate, defined to be positive in the downward direction in a 3D coordinate frame $\mathbf{r} = (\mathbf{x}, z)$, where $\mathbf{x} = (x_1, x_2)$ denotes the horizontal coordinate. The surface is illuminated by a general incident field $u^i(\mathbf{r})$, expressed as a sum of plane waves

$$u^i(\mathbf{r}) = \left(\frac{1}{2\pi}\right)^2 \int \tilde{u}^i(\mathbf{k}_x) e^{i\mathbf{k}_x \cdot \mathbf{x} - ik_z z} d^2 k_x, \quad (1)$$

where $\mathbf{k}_x = (k_{x1}, k_{x2})$, $\tilde{u}^i(\mathbf{k}_x)$ is the Fourier transform of the field at $z=0$, and $k_z = \sqrt{k - |\mathbf{k}_x|^2}$, with $\text{Im}(k_z) \geq 0$ so that the field decays at $z \rightarrow \infty$. The scattered field is given by

$$u^s(\mathbf{r}) = \left(\frac{1}{2\pi}\right)^2 \int \tilde{u}^s(\mathbf{k}_x) e^{i\mathbf{k}_x \cdot \mathbf{x} + ik_z z} d^2 k_x, \quad (2)$$

where

$$\tilde{u}^s(\mathbf{k}_x) = \int \tilde{u}^i(\mathbf{k}'_x) S(\mathbf{k}_x, \mathbf{k}'_x) d^2 k'_x, \quad (3)$$

with $S(\mathbf{k}_x, \mathbf{k}'_x)$ being the ‘‘scattering amplitude’’ (Voronovich, 1999). This definition refers to a single realization of the rough surface, but also contains the statistical characteristics of the random surface. If the roughness statistics is spatially homogeneous in the horizontal plane, then S has the form (Voronovich, 1999)

$$S(\mathbf{k}_x, \mathbf{k}'_x) = \bar{V}(\mathbf{k}_x) \delta(\mathbf{k}_x - \mathbf{k}'_x) + \Delta S(\mathbf{k}_x, \mathbf{k}'_x), \quad (4)$$

where $\bar{V}(\mathbf{k}_x)$ is the ‘‘mean reflection coefficient’’ while ΔS represents the fluctuation in the scattering amplitude with zero ensemble average. Its statistical moments satisfy

$$\begin{aligned} \overline{\Delta S} &= 0, \quad \overline{\Delta S(\mathbf{k}_{x1}, \mathbf{k}'_{x1}) \Delta S^*(\mathbf{k}_{x2}, \mathbf{k}'_{x2})} \\ &= \Delta E(\mathbf{k}_{av}, \mathbf{k}'_{av}; \delta_k) \delta(\delta_k - \delta'_k), \end{aligned} \quad (5)$$

where $\mathbf{k}_{av} = (\mathbf{k}_{x1} + \mathbf{k}_{x2})/2$, $\mathbf{k}'_{av} = (\mathbf{k}'_{x1} + \mathbf{k}'_{x2})/2$, $\delta_k = (\mathbf{k}_{x2} - \mathbf{k}_{x1})$, and $\delta'_k = (\mathbf{k}'_{x2} - \mathbf{k}'_{x1})$. The special case $\delta_k = \delta'_k = 0$ defines the ‘‘scattering cross section’’

$$\Delta E(\mathbf{k}_x, \mathbf{k}'_x; 0) = \sigma(\mathbf{k}_x, \mathbf{k}'_x). \quad (6)$$

B. Field observables: Coherent and noncoherent fields

The physical observables are the average field and the average intensity. The former, also called ‘‘the coherent field,’’ is defined by

$$u_c^s(\mathbf{r}) = \bar{u}^s(\mathbf{r}), \quad (7)$$

where the overbar denotes an average over an ensemble of realizations, and u_c^s can be described by the spectral integral (2) where, from (4), the spectral amplitude is given by

$$\tilde{u}_c^s(\mathbf{k}_x) = \tilde{u}^i(\mathbf{k}_x) \bar{V}(\mathbf{k}_x). \quad (8)$$

The scattered intensity is defined by $I^s(\mathbf{r}) = u^s(\mathbf{r}) u^{s*}(\mathbf{r})$. Averaging the measured intensity over the ensemble results yields

$$\bar{I}^s(\mathbf{r}) = I_c^s(\mathbf{r}) + I_{nc}^s(\mathbf{r}), \quad (9)$$

where $I_c^s = u_c^s u_c^{s*}$ is the coherent intensity; hence, I_{nc}^s is defined as the ‘‘noncoherent intensity.’’ It is a measure of the field fluctuations above the average (or the coherent) field. It can be expressed in terms of the second moment of S in (5), but the general expression is quite cumbersome. A simple relation is obtained in the far-field zone [see (12b)].

1. Far-zone observables

Of particular importance in various applications and in analysis are the far-zone observables. By an asymptotic evaluation of (2) for $|\mathbf{r}| \equiv r \rightarrow \infty$, one obtains

$$\begin{aligned} u^{\text{FF}}(\mathbf{r}) &= -2ik \cos \theta \frac{e^{ikr}}{4\pi r} \\ &\times \tilde{u}^s(\mathbf{k}_x) \Big|_{\mathbf{k}_x = k\mathbf{x}/r = k \cos \theta (\cos \phi, \sin \theta)}, \end{aligned} \quad (10)$$

where (θ, ϕ) are the conventional spherical angles; hence, the intensity is given by

$$I^{\text{FF}}(\mathbf{r}) = \frac{k^2 \cos^2 \theta}{4 \pi^2 r^2} \left| \tilde{u}^s(\mathbf{k}_x) \tilde{u}^{s*}(\mathbf{k}_x) \right|_{\mathbf{k}_x = k\mathbf{x}/r}. \quad (11)$$

Inserting (3) and averaging over the ensemble as in (9) results in the following expressions for the far-field coherent and noncoherent intensities:

$$I_c^{\text{FF}}(\mathbf{r}) = \frac{k^2 \cos^2 \theta}{4 \pi^2 r^2} |\bar{u}^i(\mathbf{k}_x)|^2 |\bar{V}(\mathbf{k}_x)|_{\mathbf{k}_x = k\mathbf{x}/r}^2 \quad (12a)$$

$$I_{\text{nc}}^{\text{FF}}(\mathbf{r}) = \frac{k^2 \cos^2 \theta}{4 \pi^2 r^2} \int |\bar{u}^i(\mathbf{k}'_x)|^2 \sigma(\mathbf{k}_x, \mathbf{k}'_x) d^2 k'_x |_{\mathbf{k}_x = k\mathbf{x}/r}. \quad (12b)$$

2. Noncoherent field

The noncoherent intensity $I_{\text{nc}}^s(\mathbf{r})$ is a measure of the field fluctuations above the coherent (i.e., average) field. To propagate this constituent we define the noncoherent field $u_{\text{nc}}^s(\mathbf{r})$, which provides the mathematical way to calculate the propagation of the observable $I_{\text{nc}}^s = |u_{\text{nc}}^s|^2$. In other words, u_{nc}^s can be regarded as a “deterministic realization” of a field that gives rise to the stochastic observable I_{nc}^s and particularly $I_{\text{nc}}^{\text{FF}}$.

One way to define u_{nc}^s is to construct its spectral amplitude \tilde{u}_{nc}^s from the far-field observable $I_{\text{nc}}^{\text{FF}}$ via (11), and then to propagate it via (2). The GBS representation with the GB2GB scattering matrices, introduced later on in this paper, provides an alternative to the spectral approach that can be used to propagate the stochastic observables in inhomogeneous medium.

Referring to the noncoherent part in (11), we define the spectral amplitude \tilde{u}_{nc}^s as

$$\begin{aligned} |\tilde{u}_{\text{nc}}^s(\mathbf{k}_x)| &= \frac{2 \pi r}{k \cos \theta} \sqrt{I_{\text{nc}}^{\text{FF}}(\mathbf{r})|_{\mathbf{x}/r = \mathbf{k}_x/k}} \\ &= \sqrt{\int |\bar{u}^i(\mathbf{k}'_x)|^2 \sigma(\mathbf{k}_x, \mathbf{k}'_x) d^2 k'_x}, \end{aligned} \quad (13)$$

where the second form follows from (12b).

By definition, the phase of $\tilde{u}_{\text{nc}}^s(\mathbf{k}_x)$ cannot be recovered. Yet, $\tilde{u}_{\text{nc}}^s(\mathbf{k}_x)$ contains a deterministic phase term that describes the localization of the incident wave. We recall that scattering cross section $\sigma(\mathbf{k}_x, \mathbf{k}'_x)$ is found from the far-field intensity measured with respect to a certain origin. Defining the center of mass of the incident field

$$\mathbf{x}^i = \|\mathbf{x} u^i(\mathbf{x})\| / \|u^i(\mathbf{x})\|, \quad (14)$$

where $\|\cdot\|$ represents the conventional L_2 norm, the noncoherent spectral amplitude becomes

$$\tilde{u}_{\text{nc}}^s(\mathbf{k}_x) = e^{-i \mathbf{k}_x \cdot \mathbf{x}^i} \sqrt{\int |\bar{u}^i(\mathbf{k}'_x)|^2 \sigma(\mathbf{k}_x, \mathbf{k}'_x) d^2 k'_x}. \quad (15)$$

The phase in (15) should be understood in the context of the noncoherent scattered field. It is a deterministic parameter due to the incident field localization, and not “phase reconstruction” of \tilde{u}_{nc}^s as such phase is undefined.

III. FRAME-BASED GAUSSIAN BEAM SUMMATION (GBS) METHOD

The frame-based GBS method provides an alternative to the conventional plane-wave representation. In this section we briefly review the basic concepts of this formulation, while in Sec. IV we shall use it to formulate a local spectrum representation for rough surface scattering, as an alternative to the spectral formulation of Sec. II.

We start in Sec. III A with a brief discussion on the windowed Fourier transform (WFT) frames that provide the mathematical framework for the GBS method to be presented in Sec. III B.

A. Signal representation using windowed Fourier transform (WFT) frame

For simplicity, we consider functions of one space coordinate (only in Sec. III A). The formulation is based on the construction of two dual frame sets in $L_2(\mathbb{R})$

$$\psi_{\boldsymbol{\mu}}(x) = \psi(x - m\bar{x}) e^{i n \bar{k}_x (x - m\bar{x})}, \quad (16)$$

$$\varphi_{\boldsymbol{\mu}}(x) = \varphi(x - m\bar{x}) e^{i \bar{k}_x (x - m\bar{x})}, \quad (17)$$

where $\psi(x)$ is a proper window function, $\varphi(x)$ is its dual (see below), and $\boldsymbol{\mu} = (m, n) \in \mathbb{Z}^2$ is an index. The parameters (\bar{x}, \bar{k}_x) define spatial and spectral displacement units in the continuous phase-space (x, k_x) , so that $(m\bar{x}, n\bar{k}_x)$ constitutes a discrete phase-space grid that defines the phase-space centers of the elements in $\{\psi(x)\}$ and $\{\varphi(x)\}$. For these sets to be frames, it is necessary that the unit-cell area satisfies the condition

$$\bar{x} \bar{k}_x = 2 \pi \nu, \quad \text{where } \nu < 1. \quad (18)$$

The parameter ν is the overcompleteness (or oversampling) factor of the frame. The frame is overcomplete for $\nu < 1$ and it is complete in the Gabor limit $\nu \uparrow 1$.

It is desired (though not necessary) that the windows provide a balance spatial and spectral coverage of the unit cell. This is achieved if ψ is “matched” to the grid so that

$$\Delta_x / \bar{x} = \Delta_{k_x} / \bar{k}_x, \quad (19)$$

where $\Delta_x = \|x \psi(x)\| / \|\psi(x)\|$ and $\Delta_{k_x} = \|k_x \tilde{\psi}(k_x)\| / \|\tilde{\psi}(k_x)\|$ are the spatial and spectral widths of ψ , respectively.

In general, the dual window φ in (17) needs to be calculated numerically for a given ψ and grid (\bar{x}, \bar{k}_x) , but for small ν it can be approximated analytically as

$$\varphi(x) \simeq C_0 \psi(x), \quad \text{where } C_0 = \nu / \|\psi\|^2. \quad (20)$$

A given function $u(x) \in L_2(\mathbb{R})$ can be expanded now as

$$u(x) = \sum_{\boldsymbol{\mu}} a_{\boldsymbol{\mu}} \psi_{\boldsymbol{\mu}}(x), \quad (21)$$

$$a_{\boldsymbol{\mu}} = \langle u(x), \varphi_{\boldsymbol{\mu}}(x) \rangle, \quad (22)$$

where $\langle f(x), g(x) \rangle = \int dx f(x) g^*(x)$ is the conventional L_2 inner product. In view of (17), $a_{\boldsymbol{\mu}}$ are readily identified as WFT of $u(x)$ with respect to the window $\varphi(x)$, centered about the phase-space lattice point $(m\bar{x}, n\bar{k}_x)$. Referring to

Eqs. (21)–(22), $\{\varphi_{\mu}(x)\}$ and $\{\psi_{\mu}(x)\}$ are also termed the “analysis” and “synthesis” frames, respectively.

As an example, consider a Gaussian window of the form

$$\psi(x) = e^{-x^2/2\sigma^2}, \quad \sigma^2 > 0. \quad (23)$$

From (19), this window is matched to the grid if $\sigma^2 = \bar{x}/\bar{k}_x$. From (20), $C_0 = \nu/\sqrt{\pi}\sigma$ giving, for $\nu < 0.5$

$$\varphi(x) \approx \frac{\nu}{\sqrt{\pi}\sigma} e^{-x^2/2\sigma^2}. \quad (24)$$

B. Phase-space Gaussian beam expansion

The WFT frame expansion can be used to describe the propagation of a given wave field $u(\mathbf{x}, z=0) \equiv u_0(\mathbf{x})$, which is defined on the $z=0$ plane in the three-dimensional coordinate space $\mathbf{r} = (x, z)$. To this end, we expand the data $u_0(\mathbf{x})$ using the WFT frame theory which is extended now to two space dimensions. Using properly chosen Gaussian windows, the synthesis elements $\psi_{\mu}(\mathbf{x})$ give rise to Gaussian beam (GB) propagators $B_{\mu}(\mathbf{r})$ that are used to represent the propagated field.

The 2D expansion of the WFT frame using a Gaussian window is obtained by a simple Cartesian multiplication of the 1D frame expansion. The phase-space lattice is

$$\begin{aligned} \mathbf{x}_m &= (m_1 \bar{x}_1, m_2 \bar{x}_2), \quad \mathbf{k}_{x\mathbf{n}} = (n_1 \bar{k}_{x1}, n_2 \bar{k}_{x2}), \\ \boldsymbol{\mu} = (\mathbf{m}, \mathbf{n}) &= (m_1, m_2, n_1, n_2), \end{aligned} \quad (25)$$

where \bar{x}_j and \bar{k}_{xj} , $j=1,2$ are the translation steps in the spatial and spectral domains along the x_j axes. Note that the x_1 and x_2 lattices need not be identical, but both should satisfy (18), i.e., $\bar{x}_j \bar{k}_{xj} = 2\pi\nu_j$, with $\nu_j < 1$. To be specific, we choose a square lattice with $\bar{x}_1 = \bar{x}_2 = \bar{x}$ and $\bar{k}_{x1} = \bar{k}_{x2} = \bar{k}_x$.

The WFT frame expansion of the 2D data field $u_0(\mathbf{x})$ is

$$u_0(\mathbf{x}) = \sum_{\boldsymbol{\mu}} a_{\boldsymbol{\mu}} \psi_{\boldsymbol{\mu}}(\mathbf{x}), \quad \psi_{\boldsymbol{\mu}}(\mathbf{x}) = \psi(\mathbf{x} - \mathbf{x}_m) e^{i\mathbf{k}_{x\mathbf{n}} \cdot (\mathbf{x} - \mathbf{x}_m)}. \quad (26)$$

The expansion coefficients $a_{\boldsymbol{\mu}}$ are calculated by projecting $u_0(\mathbf{x})$ onto the dual frame via [see also (50)]

$$a_{\boldsymbol{\mu}} = \langle u_0(\mathbf{x}), \varphi_{\boldsymbol{\mu}}(\mathbf{x}) \rangle. \quad (27)$$

Having expanded the field $u_0(\mathbf{x})$ in the $z=0$ plane to Gaussian frames, we propagate it into the lower homogeneous half-space $z > 0$, giving

$$u(\mathbf{r}) = \sum_{\boldsymbol{\mu}} a_{\boldsymbol{\mu}} B_{\boldsymbol{\mu}}(\mathbf{r}), \quad (28)$$

where the propagator $B_{\boldsymbol{\mu}}(\mathbf{r})$ is the radiated field due to the field distribution $\psi_{\boldsymbol{\mu}}(\mathbf{k}_x)$ on the $z=0$ plane. It is given by either one of the following alternatives:

$$B_{\boldsymbol{\mu}}(\mathbf{r}) = 2 \int d^2x' \psi_{\boldsymbol{\mu}}(\mathbf{x}') \partial_z G(\mathbf{r}|\mathbf{r}')|_{z'=0} \quad (29a)$$

$$= \left(\frac{1}{2\pi} \right)^2 \int d^2k_x \tilde{\psi}_{\boldsymbol{\mu}}(\mathbf{k}_x) e^{i\mathbf{k}_x \cdot \mathbf{x} + ik_x z}. \quad (29b)$$

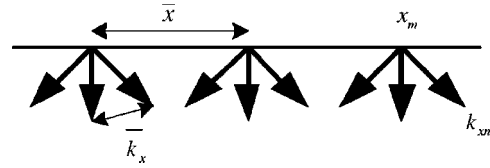


FIG. 2. The discrete phase-space beam summation representation. The beams emerge from the lattice points \mathbf{x}_m in the spectral directions $\mathbf{k}_{x\mathbf{n}}$. (\bar{x}, \bar{k}_x) are the spatial and spectral unit cell dimensions, respectively.

Expression (29a) is the Kirchhoff representation, with $G(\mathbf{r}|\mathbf{r}') = \exp\{ik|\mathbf{r} - \mathbf{r}'|\}/4\pi|\mathbf{r} - \mathbf{r}'|$ being the free-space Green's function and ∂_z denotes the derivative with respect to z' at the source point \mathbf{r}' . Expression (29b) is the plane-wave representation, with $\tilde{\psi}_{\boldsymbol{\mu}}(\mathbf{k}_x) = \tilde{\psi}(\mathbf{k}_x - \mathbf{k}_{x\mathbf{n}}) e^{-i\mathbf{k}_x \cdot \mathbf{x}_m}$ being the spectral counterpart of $\psi_{\boldsymbol{\mu}}(\mathbf{x})$.

Since $\psi_{\boldsymbol{\mu}}$ are localized about $(\mathbf{x}_m, \mathbf{k}_{x\mathbf{n}})$ in the phase space, $B_{\boldsymbol{\mu}}$ describes a beam field that emerges from \mathbf{x}_m in the $z=0$ plane in the direction (unit vectors are denoted by a caret)

$$\hat{\mathbf{k}}_{\mathbf{n}} = (\mathbf{k}_{x\mathbf{n}}, k_{z\mathbf{n}})/k = (\cos \phi_{\mathbf{n}} \sin \theta_{\mathbf{n}}, \sin \phi_{\mathbf{n}} \sin \theta_{\mathbf{n}}, \cos \theta_{\mathbf{n}}). \quad (30)$$

Thus, propagating beams are obtained only for small \mathbf{n} such that $|\mathbf{k}_{x\mathbf{n}}| < k$: For $|\mathbf{k}_{x\mathbf{n}}| > k$ the field of $B_{\boldsymbol{\mu}}$ decays away from the $z=0$ plane. In practice we ignore all the evanescent beams and use in (28) only the beams with $|\mathbf{k}_{x\mathbf{n}}| < 1$, thereby expressing the radiated field as a discrete phase-space superposition of beams that emerge from all lattice points \mathbf{x}_m in the aperture plane and in all real lattice directions $\mathbf{k}_{x\mathbf{n}}$; Fig. 2.

It should be noted that for a given observation point \mathbf{r} , we truncate the summation in (28) so that only those beam propagators that pass near \mathbf{r} are included [actually we include those that pass within a 3-beamwidth distance from \mathbf{r} (Steinberg *et al.*, 1991; Shlivinski *et al.*, 2004)]. Viewed from a phase-space perspective, this localizes the summation to the vicinity of the so-called *geometrical observation manifold* that defines the initiation points and directions of the beams whose axis pass through the observation point. In free space, this manifold is defined by

$$(\mathbf{x} - \mathbf{x}_m)/z \approx \mathbf{k}_{x\mathbf{n}}/k_{z\mathbf{n}}. \quad (31)$$

This truncation greatly reduces the computation cost.

Next, we consider the Gaussian windows of (23), which will be rewritten henceforth in the form

$$\begin{aligned} \psi(x) &= e^{-k|x|^2/2b}, \quad \text{and} \\ \varphi(x) &\approx \nu^2(k/\pi b) e^{-k|x|^2/2b}, \quad b = k\bar{x}/\bar{k}_x, \end{aligned} \quad (32)$$

where, for analytic simplicity, we use $\nu < 0.5$ so that φ can be approximated as in (24), while the condition on b follows from (19). We also use $kb \gg 1$ so that ψ is wide on a wavelength scale and it yields collimated Gaussian beams (GB) propagators. Analytic expressions for $B_{\boldsymbol{\mu}}$ can be obtained by evaluating the integral (29b) asymptotically via saddle-point integration as detailed in Melamed (1997). One finds that $B_{\boldsymbol{\mu}}$ is an astigmatic GB that can be described in the most physically appealing format by utilizing the beam coordinates $\mathbf{r}_b = (x_{b1}, x_{b2}, z_b)$ defined for a given phase-space point $\boldsymbol{\mu}$. The

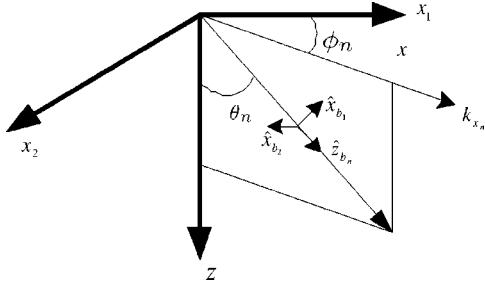


FIG. 3. The beam coordinate system (33)–(34) corresponding to a given beam propagator B_μ . Here, \hat{z}_{b_n} is the beam axis and $\hat{x}_{b_{1,2}}$ are the transverse axes as defined there.

beam axis z_b emerges from $\mathbf{r}_m = (\mathbf{x}_m, z=0)$ in the direction $\hat{\mathbf{k}}_n$ of (30). The transverse coordinates $\mathbf{x}_b = (x_{b_1}, x_{b_2})$ are chosen such that the projection of the unit vector \hat{x}_{b_1} on the $z=0$ plane coincides with the direction of \mathbf{k}_{x_n} , while $\hat{x}_{b_2} = \hat{z}_b \times \hat{x}_{b_1}$. With this choice, the linear phase $\mathbf{k}_{x_n} \cdot \mathbf{x}$ implied by the window function ψ_μ [see (26)] is operative in the x_{b_1} direction *but not* in the x_{b_2} direction. The beam and physical coordinates are related by (see Fig. 3)

$$\mathbf{r}_b = \mathbf{\Omega}_n(\mathbf{r} - \mathbf{r}_m), \quad (33)$$

where the unitary rotation matrix $\mathbf{\Omega}_n$ is defined so that the transverse coordinates are rotated about the z_b axis such that $\mathbf{x}_{b_2} \cdot \mathbf{k}_{x_n} = 0$

$$\mathbf{\Omega}_n = \begin{pmatrix} \cos \theta_n \cos \phi_n & \cos \theta_n \sin \phi_n & -\sin \theta_n \\ -\sin \phi_n & \cos \phi_n & 0 \\ \sin \theta_n \cos \phi_n & \sin \theta_n \sin \phi_n & \cos \theta_n \end{pmatrix}. \quad (34)$$

Along this axis, the astigmatic GB field can be expressed as

$$B_\mu(\mathbf{r}) \approx \left(\frac{\det \Gamma(z_b)}{\det \Gamma(0)} \right)^{1/2} e^{ikz_b} e^{i(k/2)\mathbf{x}_b^T \Gamma(z_b) \mathbf{x}_b}, \quad (35)$$

where $\Gamma(z_b)$ is a 2×2 complex symmetric matrix with $\text{Im} \Gamma(z_b)$ positive definite. As a result, the imaginary part of the quadratic form $\mathbf{x}_b^T \Gamma(z_b) \mathbf{x}_b = x_{b_1}^2 \Gamma_{11} + 2x_{b_1} x_{b_2} \Gamma_{12} + x_{b_2}^2 \Gamma_{22}$ increases away from the z_b axis, leading to the Gaussian decay of the beam field. In free space it is found from the initial condition $\Gamma(0)$ via

$$\Gamma(z_b) = [z_b \mathbf{I} + \Gamma^{-1}(0)]^{-1}, \quad (36)$$

where \mathbf{I} is the unit matrix. One observes that indeed $\text{Im} \Gamma(z_b)$ is positive definite for all z_b provided that $\text{Im} \Gamma(0)$ is positive definite. In inhomogeneous medium, $\Gamma(z_b)$ is calculated by solving a proper differential equation (the so-called “dynamic ray-tracing equation”) along the ray trajectory that defines the beam axis (Červený *et al.*, 1982; Babich and Popov, 1989).

In the coordinate system defined above, $\Gamma(0)$, and thereby $\Gamma(z_b)$, is diagonal and is given by

$$\Gamma_{\ell,\ell}(0) = i/F_\ell, \quad F_1 = b \cos^2 \theta_n, \quad F_2 = b, \quad \ell = 1, 2. \quad (37)$$

Substituting in (35)–(36), the GB field in that system is given, explicitly, by

$$B_\mu(\mathbf{r}) \approx \sqrt{\frac{-iF_1}{z_b - iF_1} \frac{-iF_2}{z_b - iF_2}} \times \exp \left\{ ik \left[z_b + \frac{x_{b_1}^2/2}{z_b - iF_1} + \frac{x_{b_2}^2/2}{z_b - iF_2} \right] \right\}. \quad (38)$$

The beam characteristics in the x_{b_ℓ} cuts are found by separating the corresponding phase term in (38) into real and imaginary parts in the form

$$ik \frac{x_{b_\ell}^2/2}{z_b - iF_\ell} = ik \frac{x_{b_\ell}^2/2}{R_\ell(z_b)} - \frac{x_{b_\ell}^2/2}{W_\ell^2(z_b)}, \quad (39)$$

giving

$$W_\ell(z_b) = W_{0_\ell} \sqrt{1 + (z_b/F_\ell)^2}, \quad W_{0_\ell} = \sqrt{F_\ell/k}, \quad (40)$$

$$R_\ell = z_b + F_\ell^2/z_b.$$

Thus, $W_\ell(z_b)$ is the beamwidth, W_{0_ℓ} is the waist, $R_\ell(z_b)$ is the wavefront radius of curvature. F_ℓ is readily identified as the diffraction length in the x_{b_ℓ} cut. For $z_b \ll F_\ell$ we have $W_\ell(z_b) \approx W_{0_\ell}$, while for $z_b \gg F_\ell$ the beam opens up along the constant diffraction angles $\Theta_\ell = 1/\sqrt{F_\ell k}$.

The GB model (35)–(40) is an asymptotic approximation to the exact propagators as defined in (29). The validity increases with the beam collimation kb , but it deteriorates for large inclination angles $\theta_n = 0$ (see Shlivinski *et al.*, 2004, Fig. 7 for a detailed study of the error as a function of these parameters). In the numerical simulations of Sec. VC we used $kb = 5000$; hence, the propagator error for $\theta_n = 60^\circ$ is less than 1%.

IV. GAUSSIAN BEAM SUMMATION (GBS) ANALYSIS OF ROUGH SURFACE SCATTERING

We now use the GBS theory from the previous section for phase-space (local spectrum) analysis of the rough surface scattering phenomenology discussed in Sec. II. The scattered field will be analyzed using the frame $\{\psi_\mu\}$ with respect to the $z=0$ plane (as discussed in Sec. III).

A. GB modeling of the incident field

We shall consider two options to describe the incident field $u^i(\mathbf{r})$ that hits the rough surface from below (i.e., from $z > 0$). The first is to expand u^i using the same frame $\{\psi_\mu\}$ which is used to expand the scattered field. The other is to describe u^i using a GBS with reference to a different frame and a different reference plane.

An example for the latter approach is when the field of a source located below the rough surface is expanded using a frame $\{\psi_\mu^{(i)}\}$ with respect to the source aperture plane (the aperture can be horizontal, vertical, or at any other convenient angle) or with respect to a point source; see Fig. 1. In any case, the incident field hitting the upper surface at $z=0$ has the general form [see (28)]

$$u^i(\mathbf{r}) = \sum_{\mu^i} a_{\mu^i} B_{\mu^i}^i(\mathbf{r}), \quad (41)$$

where $B_{\mu^i}^i(\mathbf{r})$ are GBs of the form (38) emerging from the source domain toward the rough surface, μ^i is an index, and a_{μ^i} are the expansion coefficients.

Another example for the second approach appears in Gordon *et al.*, 2005, where we consider wave propagation between two rough surfaces and use, in general, different frames $\{\psi_{\mu_1}^{(1)}\}$ and $\{\psi_{\mu_2}^{(2)}\}$ for expansion in the upper and lower surfaces, respectively. In that case, the incident field hitting the upper surface at $z=0$ is expanded in terms of the GBs $B_{\mu_2}^{(2)}(\mathbf{r})$ that emerge from the lower surface and hit the upper surface, i.e., it has the general form in (41) except that the superscript i is now replaced by (2).

Henceforth, we shall consider the second approach discussed in (41) above, since the first approach, namely to expand u^i using $\{\psi_{\mu^i}\}$, can be treated as a special simplified case of the second approach.

B. GB-to-GB scattering matrices

Without loss of generality, we shall assume below that the incident field is described by a single element in (41) and has the typical GB form in (35)

$$u^i(\mathbf{r}) = B^i(\mathbf{r}) = A^i(z_b^i) e^{ik[z_b^i + (1/2)\mathbf{x}_b^{iT} \mathbf{\Gamma}^i(z_b^i) \mathbf{x}_b^i]}, \quad (42)$$

where (\mathbf{x}_b^i, z_b^i) are the beam coordinates with respect to the axis of this beam, $A(z_b^i)$ is the slowly varying amplitude of this beam along the z_b^i axis, and $\mathbf{\Gamma}^i(z_b^i)$ is a complex symmetric matrix with $\text{Im } \mathbf{\Gamma}^i$ positive definite as discussed in (35) and (36).

At the scattering reference plane $z=0$, the incident field is identified by the following parameters:

- (i) A_0^i and $\mathbf{\Gamma}_0^i$, which are the values of A^i and $\mathbf{\Gamma}^i$ at the point where the beam axis intersects the $z=0$ plane. We use these values to describe the incident GB even at off-axis points on the plane $z=0$ plane (clearly this is an approximation since $\mathbf{\Gamma}^i$ are functions of z_b^i so that for tilted beams they are not constant on the $z=0$ plane, yet it is legitimate for narrow beams).
- (ii) The phase-space center of the beam at $z=0$: $\mathbf{x}^i = (x_1^i, x_2^i)$ and $\mathbf{k}_x^i = (k_{x1}^i, k_{x2}^i) = k \sin \theta^i (\cos \phi^i, \sin \phi^i)$ where (θ^i, ϕ^i) , with $\theta^i > \pi/2$, define the beam direction.
- (iii) Without loss of generality we use the coordinate convention in (33) and choose the beam coordinate system $\mathbf{x}_b^i = (x_{b1}^i, x_{b2}^i)$ in (42) such that the projection of $\hat{\mathbf{x}}_{b1}^i$ on the $z=0$ plane is along the \mathbf{k}_x^i axis, while $\hat{\mathbf{x}}_{b2}^i = \hat{z}_b^i \times \hat{\mathbf{x}}_{b1}^i$ (in other words, $\hat{\mathbf{x}}_{b2}^i \cdot \mathbf{k}_x^i = 0$) (compare Fig. 3).

In order to evaluate the scattered field as in (2)–(3), we need to calculate the plane-wave spectrum of the incident field via

$$\tilde{B}^i(\mathbf{k}_x) = \int B^i(\mathbf{x}, z=0) e^{i\mathbf{k}_x \cdot \mathbf{x}} d^2x. \quad (43)$$

To this end, we need to express the incident beam field in terms of the physical (or geographical) coordinates \mathbf{x} at $z=0$.

Recalling the coordinate definition above, we note that \mathbf{x}_b^i are related to the geographical coordinates via [cf. (33)]

$$\mathbf{x}_b^i = \boldsymbol{\zeta}^i (\mathbf{x} - \mathbf{x}^i), \quad (44)$$

where the 2×2 matrix $\boldsymbol{\zeta}^i$ is defined via [cf. (33)]

$$\begin{aligned} \boldsymbol{\Omega}^i &= \begin{pmatrix} \begin{bmatrix} -\cos \theta^i \cos \phi^i & -\cos \theta^i \sin \phi^i \\ \sin \phi^i & -\cos \phi^i \end{bmatrix} & \begin{bmatrix} \sin \theta^i \\ 0 \end{bmatrix} \\ \begin{bmatrix} \sin \theta^i \cos \phi^i & \sin \theta^i \sin \phi^i \end{bmatrix} & \cos \theta^i \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{\zeta}^i & \boldsymbol{\eta}_1^i \\ \boldsymbol{\eta}_2^i & \boldsymbol{\epsilon}^i \end{pmatrix}. \end{aligned} \quad (45)$$

Note that $\boldsymbol{\Omega}^i$ is slightly different from $\boldsymbol{\Omega}_n$ (34) due to the difference between an incoming and an outgoing direction of the beam, respectively. Using these definitions

$$B^i(\mathbf{x}, z=0) = A_0^i e^{(ik/2)(\mathbf{x} - \mathbf{x}^i)^T \boldsymbol{\zeta}^i \mathbf{\Gamma}_0^i \boldsymbol{\zeta}^i (\mathbf{x} - \mathbf{x}^i) + i\mathbf{k}_x^i \cdot \mathbf{x}}. \quad (46)$$

Inserting (46) into (43) and evaluating the resulting integral via saddle-point integration, which, in this case, coincides with the exact solution, yields

$$\tilde{B}^i(\mathbf{k}_x) = A_0^i \frac{2\pi i}{k \sqrt{\det(\mathbf{\Gamma}_k^i)^{-1}}} e^{(-i/2k)(\mathbf{k}_x - \mathbf{k}_x^i)^T \mathbf{\Gamma}_k^i (\mathbf{k}_x - \mathbf{k}_x^i) - i(\mathbf{k}_x - \mathbf{k}_x^i) \cdot \mathbf{x}^i}, \quad (47)$$

where

$$\mathbf{\Gamma}_k^i = (\boldsymbol{\zeta}^{iT} \mathbf{\Gamma}_0^i \boldsymbol{\zeta}^i)^{-1}, \quad (48)$$

with the subscript k (or x , see below) indicating that the matrix relates to \mathbf{k}_x (or \mathbf{x}), and we take the root of $\sqrt{\det(\mathbf{\Gamma}_k^i)^{-1}}$, which has a positive imaginary part (recall that $\text{Im } \mathbf{\Gamma}_0^i$ is positive definite, so that $\text{Im } \mathbf{\Gamma}_k^i$ is negative definite; hence, this definition of the square root provides a continuous function of $\text{Re } \mathbf{\Gamma}_k^i$).

Finally, the plane-wave representation of the coherent and incoherent scattered fields is given by (8) and (15), with $\tilde{u}^i(\mathbf{k}_x)$ replaced by $\tilde{B}^i(\mathbf{k}_x)$ of (47).

Here, however, we shall describe the coherent and incoherent scattered fields by using the frame expansion of Sec. III B. Thus, at the $z=0$ plane, the fields are given by

$$u_c^s(\mathbf{x}, z=0) = \sum_{\mu} S_c(B^i, \psi_{\mu}) \psi_{\mu}(\mathbf{x}), \quad (49a)$$

$$u_{nc}^s(\mathbf{x}, z=0) = \sum_{\mu} S_{nc}(B^i, \psi_{\mu}) \psi_{\mu}(\mathbf{x}). \quad (49b)$$

The expansion coefficients, the GB-to-GB (GB2GB) scattering matrices, are obtained by projecting B^i onto the dual frame functions as in (27). Instead of (27), however, we shall use its spectral counterpart

$$a_{\mu} = \left(\frac{1}{2\pi} \right)^2 \langle \tilde{u}_0(\mathbf{k}_x), \tilde{\varphi}_{\mu}(\mathbf{k}_x) \rangle. \quad (50)$$

Thus, using (8) and (15) with $\tilde{B}^i(\mathbf{k}_x)$ of (47), the scattering matrices are calculated via

$$S_c(B^i, \psi_\mu) = \left(\frac{1}{2\pi}\right)^2 \int \bar{B}^i(\mathbf{k}_x) \bar{V}(\mathbf{k}_x) \bar{\varphi}_\mu^*(\mathbf{k}_x) d^2k_x, \quad (51a)$$

$$S_{nc}(B^i, \psi_\mu) = \left(\frac{1}{2\pi}\right)^2 \times \int e^{-i\mathbf{k}_x \cdot \mathbf{x}^i} \sqrt{|\sigma(\mathbf{k}_x, \mathbf{k}'_x)| \bar{B}^i(\mathbf{k}'_x)|^2 d^2k'_x} \times \bar{\varphi}_\mu^*(\mathbf{k}_x) d^2k_x, \quad (51b)$$

where $\bar{\varphi}_\mu(\mathbf{k}_x) = \bar{\varphi}(\mathbf{k}_x - \mathbf{k}_{xn}) e^{-i\mathbf{k}_x \cdot \mathbf{x}_m}$ is the spectral counterpart of $\varphi_\mu(\mathbf{x})$. For the Gaussian mother windows in (32) which are used here, we have

$$\bar{\varphi}_\mu(\mathbf{k}_x) = \nu^2 \frac{k}{\pi b} \tilde{\psi}_\mu(\mathbf{k}_x) = 2\nu^2 e^{-b|\mathbf{k}_x - \mathbf{k}_{xn}|^2/2k - i\mathbf{k}_x \cdot \mathbf{x}_m}. \quad (52)$$

Finally, the field representation in (49) is propagated using the GBS representation of (41), giving GBS representation for the scattered coherent and noncoherent fields

$$u_c^s(\mathbf{r}) = \sum_\mu S_c(B^i, \psi_\mu) B_\mu(\mathbf{r}), \quad (53a)$$

$$u_{nc}^s(\mathbf{r}) = \sum_\mu S_{nc}(B^i, \psi_\mu) B_\mu(\mathbf{r}). \quad (53b)$$

An analytical example of a weakly rough surface, for which the scattering matrices can be calculated in closed form via the small perturbation approximation, is presented in Sec. V.

V. EXAMPLE: WEAKLY ROUGH SURFACE

The formulation above can be applied to any rough surface using either measured or numerically generated scattering data. Here, we shall derive analytic results for weakly rough surface with Gaussian spectrum with l_h correlation length, where the standard deviation of the roughness satisfies $k\sigma_h \ll 1$. Using the small perturbation approximation, the statistical moments are given by (Voronovich, 1999)

$$\bar{V}(\mathbf{k}_x) = -1 + 8\pi^2 \sigma_h^2 k_z^2, \quad (54)$$

$$\sigma(\mathbf{k}_x, \mathbf{k}'_x) = 4\pi \sigma_h^2 l_h^2 k_z'^2 e^{-|\mathbf{k}_x - \mathbf{k}'_x|^2 l_h^2/4}. \quad (55)$$

The coherent and noncoherent S matrices of (51) can be evaluated analytically using saddle-point integration. It is assumed here that the incident beam and the expansion beams are collimated, i.e., $kb \gg 1$. This validates the use of the saddle-point evaluation.

Below, we shall calculate the S matrices and then compare numerically the coherent and noncoherent fields obtained via the GB2GB formulation to the those calculated via the plane-wave formulation.

A. Coherent field

For the coherent field, inserting (54), (47), and (52) into (51a) results in the following S matrix:

$$S_c(B^i, \psi_\mu) = \left(\frac{1}{2\pi}\right)^2 \int \frac{A_0^i 2\pi i}{k \sqrt{\det(\Gamma_k^i)^{-1}}} \times e^{(-i/2k)(\mathbf{k}_x - \mathbf{k}'_x)^T \Gamma_k^i (\mathbf{k}_x - \mathbf{k}'_x) - i(\mathbf{k}_x - \mathbf{k}'_x) \cdot \mathbf{x}^i} \times (-1 + 8\pi^2 \sigma_h^2 k_z^2) 2\nu^2 \times e^{-b|\mathbf{k}_x - \mathbf{k}_{xn}|^2/2k - i\mathbf{k}_x \cdot \mathbf{x}_m} d^2k_x, \quad (56)$$

where Γ_k^i is given by (48) with (73), and we note that $\text{Im} \Gamma_k^i$ is negative definite.

The integral in (56) has the standard form addressed in the Appendix; its saddle-point evaluation is

$$S_c(B^i, \psi_\mu) = \frac{iA_0^i 2\nu^2}{\sqrt{\det(\Gamma_k^i)^{-1}} \sqrt{\det(\Gamma_x^c)^{-1}}} e^{i\mathbf{k}_x \cdot \mathbf{x}^i} \times (-1 + 8\pi^2 \sigma_h^2 k_z^2) \times \exp \left[-\frac{i}{2k} (\mathbf{k}_{xn} - \mathbf{k}'_x)^T \Gamma_k^c (\mathbf{k}_{xn} - \mathbf{k}'_x) + i/2k (\mathbf{x}_m - \mathbf{x}^i)^T \Gamma_x^c (\mathbf{x}_m - \mathbf{x}^i) + i\mathbf{k}^c \cdot (\mathbf{x}_m - \mathbf{x}^i) \right], \quad (57)$$

where

$$\Gamma_x^c = (\Gamma_k^i - ib\mathbf{I})^{-1}, \quad (58)$$

$$\Gamma_k^c = -ib\Gamma_x^c \Gamma_k^i = ((\Gamma_k^i)^{-1} + ib^{-1}\mathbf{I})^{-1}, \quad (59)$$

$$\mathbf{k}^c = \Gamma_x^c (\Gamma_k^i \mathbf{k}'_x - ib\mathbf{k}_{xn}), \quad (60)$$

$$\mathbf{k}_s = \mathbf{k}^c + k\Gamma_x^c (\mathbf{x}_m - \mathbf{x}^i), \quad k_{z_s} = k_z(\mathbf{k}_s), \quad (61)$$

and, referring to (58), $\text{Im}(\Gamma_x^c)^{-1}$ is negative definite; hence, we take the root of $\sqrt{\det(\Gamma_x^c)^{-1}}$ which has a negative imaginary part. Also, \mathbf{k}_s is the stationary point of the integral (56) and is required for the mean reflection coefficient.

As can be seen, the S matrix is localized in phase space, around $(\mathbf{x}_m, \mathbf{k}_{xn}) = (\mathbf{x}^i, \mathbf{k}'_x)$, with the localization parameters Γ_x^c and Γ_k^c , respectively, which are determined by the Γ^i matrix of the incident beam, and also by the collimation parameter b used for the expansion beams. This corresponds to the δ function usually associated with the spectrum of the coherent field. If we take both incident and scattered beams to be very wide (i.e., approximating plane waves) we will get complete coherent reflection in the specular direction, with no space localization. On the other hand, if we take very narrow beams that spread like a point-sources field, we get a specular scattering emerging from the point where the incident beam hits the surface. Due to the independence of the mean reflection coefficient on the roughness spectrum in the small perturbation method, the roughness determines the amplitude of the coherent S matrix and not its localization in phase space. These results are readily discerned in Fig. 5(b).

B. Noncoherent field

The noncoherent S matrix is obtained by inserting (55), (47), and (52) into (51b). We start by evaluating the inner integral in (51b) for the spectral noncoherent scattered intensity [see (13)]

$$|\bar{u}_{\text{nc}}^s(\mathbf{k}_x)|^2 = \int 4\pi\sigma_h^2 l_h^2 k_z'^2 e^{-|\mathbf{k}'_x - \mathbf{k}_x|^2 l_h^2 / 4} \left| \frac{A_0^i 2\pi i}{k \sqrt{\det(\mathbf{\Gamma}_k^i)^{-1}}} \right. \\ \left. \times e^{(-i/2k)(\mathbf{k}'_x - \mathbf{k}_x^i)^T \mathbf{\Gamma}_k^i (\mathbf{k}'_x - \mathbf{k}_x^i) - i(\mathbf{k}'_x - \mathbf{k}_x^i) \cdot \mathbf{x}^i} \right|^2 d^2 k'_x. \quad (62)$$

This integral has the same form as before, and its saddle-point evaluation is

$$|\bar{u}_{\text{nc}}^s(\mathbf{k}_x)|^2 = \frac{32\pi^4 \sigma_h^2 l_h^2 |A_0^i|^2}{|\det(\mathbf{\Gamma}_k^i)^{-1}| k} \sqrt{\det(\mathbf{\Gamma}_k^i)} \\ \times k_{z_s}^2 e^{-(\mathbf{k}_x - \mathbf{k}_x^i)^T \mathbf{\Gamma}_k^i (\mathbf{k}_x - \mathbf{k}_x^i) / k}, \quad (63)$$

where

$$\mathbf{\Gamma}_k^i = (kl_h^2/4\mathbf{I} - \text{Im} \mathbf{\Gamma}_k^i)^{-1}, \quad (64)$$

$$\mathbf{k}_s = \mathbf{\Gamma}_k^i (2 \text{Im} \mathbf{\Gamma}_k^i \mathbf{k}_x^i - \mathbf{k}_x k l_h^2 / 2). \quad (65)$$

Note that $\mathbf{\Gamma}_k^i$ is real positive.

The spectral intensity in (63) has the form of a Gaussian centered on the specular direction. The spectral localization, via the matrix $\mathbf{\Gamma}_k^i$, is affected by an interplay between the rough surface correlation length l_h and the width of the incident beam [recall from (46)–(48) that $-\text{Im} \mathbf{\Gamma}_k^i$ in (64) is proportional to kW_x^i , where W_x^i is the projected width of the incident beam on the $z=0$ plane]. Thus, if $W_x^i \gg l_h$, then $\mathbf{\Gamma}_k^i \approx kl_h^2/4\mathbf{I}$ and the spectral intensity is governed by $\sigma(\mathbf{k}_x, \mathbf{k}'_x)$, whereas if $l_h \gg W_x^i$, then $\mathbf{\Gamma}_k^i \approx -\text{Im} \mathbf{\Gamma}_k^i$ and the parameters of the rough surface vanish from the exponent.

Inserting (62) together with (52) into (51b) results in the noncoherent S matrix, again in a similar integral form

$$S_{\text{nc}}(B^i, \psi_\mu) = \left(\frac{1}{2\pi} \right)^2 \int d^2 k_x \frac{8\sqrt{2}\pi^2 \sigma_h l_h |A_0^i|}{\sqrt{k|\det(\mathbf{\Gamma}_k^i)^{-1}|}} \nu^2 \\ \times k_{z_s} [\det(\mathbf{\Gamma}_k^i)]^{1/4} e^{-(\mathbf{k}_x - \mathbf{k}_x^i)^T \mathbf{\Gamma}_k^i (\mathbf{k}_x - \mathbf{k}_x^i) / 2k - i\mathbf{k}_x \cdot \mathbf{x}^i} \\ \times e^{-b|\mathbf{k}_x - \mathbf{k}_{\text{xn}}|^2 / 2k - i\mathbf{k}_x \cdot \mathbf{x}_m}, \quad (66)$$

The saddle-point evaluation of this integral results in

$$S_{\text{nc}}(B^i, \psi_\mu) = \frac{i|A_0^i| 4\pi \nu^2 \sqrt{2k} \sigma_h l_h k_z' [\det(\mathbf{\Gamma}_k^i)]^{1/4}}{\sqrt{|\det(\mathbf{\Gamma}_k^i)^{-1}|} \sqrt{\det(\mathbf{\Gamma}_x^{\text{nc}})^{-1}}} \\ \times \exp[-(\mathbf{k}_{\text{xn}} - \mathbf{k}_x^i)^T \mathbf{\Gamma}_k^{\text{nc}} (\mathbf{k}_{\text{xn}} - \mathbf{k}_x^i) / 2k \\ - k(\mathbf{x}_m - \mathbf{x}^i)^T \mathbf{\Gamma}_x^{\text{nc}} (\mathbf{x}_m - \mathbf{x}^i) / 2 \\ + i\mathbf{k}^{\text{nc}} \cdot (\mathbf{x}_m - \mathbf{x}^i)], \quad (67)$$

where

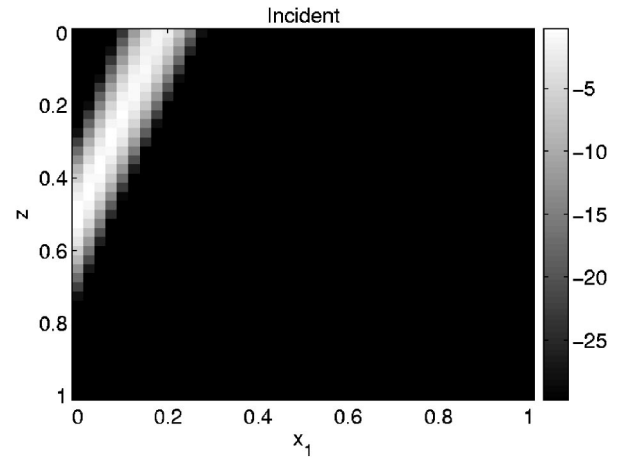


FIG. 4. Incident field intensity in dB. $b^i=1$, $k=1000$. Origin of incident field is at $(0, 0, 0.5)$ and exit angle is at $\theta=20^\circ$, $\varphi=0^\circ$. The figure is plotted in the (x_1, z) plane for $x_2=0$.

$$\mathbf{\Gamma}_x^{\text{nc}} = (\mathbf{\Gamma}_k^i + b\mathbf{I})^{-1}, \quad (68)$$

$$\mathbf{\Gamma}_k^{\text{nc}} = b\mathbf{\Gamma}_x^{\text{nc}} \mathbf{\Gamma}_k^i = ((\mathbf{\Gamma}_k^i)^{-1} + b^{-1}\mathbf{I})^{-1}, \quad (69)$$

$$\mathbf{k}_s^{\text{nc}} = \mathbf{\Gamma}_x^{\text{nc}} (\mathbf{\Gamma}_k^i \mathbf{k}_x^i + b\mathbf{k}_{\text{xn}}), \quad (70)$$

$$\mathbf{k}_s = \mathbf{k}_s^{\text{nc}} + k\mathbf{\Gamma}_x^{\text{nc}} (\mathbf{x}_m - \mathbf{x}^i), \quad (71)$$

$$\mathbf{k}'_s = \mathbf{\Gamma}_x^i \left(2 \text{Im} \mathbf{\Gamma}_k^i \mathbf{k}_x^i - \frac{kl_h^2}{2} \mathbf{k}_s \right), \quad k'_{z_s} = k_z(\mathbf{k}'_s). \quad (72)$$

Note that $\mathbf{\Gamma}_x^{\text{nc}}$ is real positive.

The noncoherent S matrix in (67) is localized in phase space around $(\mathbf{x}_m, \mathbf{k}_{\text{xn}}) = (\mathbf{x}^i, \mathbf{k}'_x)$, with the localization parameters $\mathbf{\Gamma}_x^{\text{nc}}$ and $\mathbf{\Gamma}_k^{\text{nc}}$, respectively, which are determined by the $\mathbf{\Gamma}$ matrices of the incident and scattered beams and the correlation length of the rough surface. As discussed following (63), one readily observes that for small correlation lengths compared to the incident beamwidth, S_{nc} is spectrally wide and its spatial localization is determined by the ratio between the correlation length and the scattered beam's parameter, b . For large correlation lengths compared to the incident beamwidth, the localization of S_{nc} is determined by the incident and scattered $\mathbf{\Gamma}$'s only.

C. Numerical comparison

1. Problem setup

We consider a homogeneous medium with wave number $k=1000$. The rough surface at $z=0$ is illuminated by a collimated, circular-symmetric Gaussian beam with $b^i=1$, originating from $\mathbf{r}_0=(0,0,0.5)$ with initial direction $\theta=20^\circ$, $\varphi=0^\circ$ (see Fig. 4). From (35) and (36), it has the form

$$B^i(\mathbf{r}) = \frac{-ib^i}{z_b - ib^i} \exp\left[ik \left(z_b + \frac{1}{2} \frac{|\mathbf{x}_b|^2}{z_b - ib^i} \right) \right], \quad (73)$$

where (\mathbf{x}_b, z_b^i) are the incident beams coordinates measured along the beam axis from the point \mathbf{r}_0 . The matrix $\mathbf{\Gamma}_0^i$ in (42) is then given by $\mathbf{\Gamma}_0^i = (0.5/\cos 20^\circ - ib^i)^{-1} \delta_{ij}$.

The rough surface parameters are $k\sigma_h=0.1$ for the small perturbation approximation and the correlation length is smaller than the width of the incident beam such that

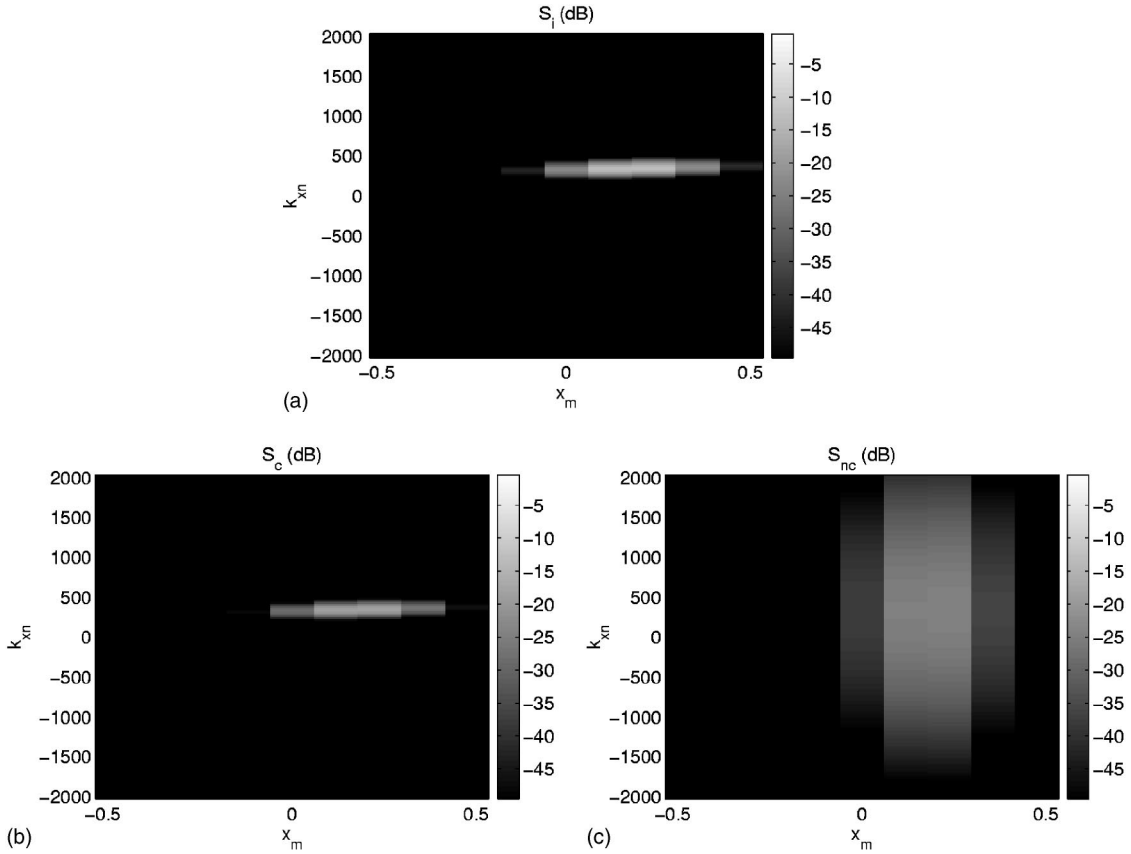


FIG. 5. GB2GB scattering matrices for the incident beam in Fig. 4, shown in the (x_1, k_{x_1}) plane where $(x_2, k_{x_2}) = (0, 0)$. Expansion parameters: $b = 5$, $\nu = 0.45$, $\bar{x} = 0.118\ 890$, $\bar{k} = 23.779\ 964$. Surface parameters: $k\sigma = 0.1$, $kl^2/b^i = 0.1$. Each point in the discrete lattice represents a single beam. (a) Phase-space coefficients of the incident field. (b) Coherent GB2GB scattering matrix. (c) Noncoherent GB2GB scattering matrix.

$l_h^2/W_0^2 = 0.1$, where $W_0 = \sqrt{b^i/k}$ [see (40)]. The expansion frame utilizes Gaussian windows of the form in (32) and (52), where we choose $b = 5$ for collimated propagators (i.e., $kb = 5000$). From (32) and (18) we find $\bar{x} = \sqrt{2\pi\nu b/k}$ and $\bar{k}_x = \sqrt{2\pi\nu k/b}$. It should be recalled that in all the field computations we include only those beams that pass near the observation point [see the geometrical observation approach discussed in connection with (31)].

2. Results

The coherent and noncoherent scattering matrices S_c and S_{nc} are shown in Figs. 5(b) and (c), respectively. In order to understand these results we also show in Fig. 5(a) the expansion coefficients $S_i(B^i, \psi_\mu)$ of the incident field with respect to the same frame $\psi_\mu(x)$. These coefficients can be calculated by setting $\bar{V}(\mathbf{k}_x) = 1$ in (51a). The plots show phase-space distributions of the matrices in the (x_1, k_{x_1}) plane for $(x_2, k_{x_2}) = (0, 0)$.

As can be seen, S_c represents essentially a specular scattering; hence, its phase-space distribution in Fig. 5(b) replicates the phase-space distribution of the incident field in Fig. 5(a) except for the multiplication term $\bar{V}(\mathbf{k}_x)$. S_{nc} , on the other hand, is still localized in space within the support of the incident field, but it is weaker and has a much wider spectral spread. These observations are supported by the analytical analysis in Secs. V A and V B.

Finally, the propagating coherent and noncoherent intensities calculated via the GBS representation are shown in Figs. 6(a) and 7(a), respectively. The coherent field is described by specular spectral reflections [see (8)] and therefore has the structure of the incident beam, whereas the noncoherent field spreads away and decays from the center point \mathbf{x}_0 of the incident field.

The high accuracy of our GBS approach to calculate the coherent and noncoherent fields is demonstrated in Figs. 6(b) and 7(b). Figure 6(b) depicts the absolute error of the GBS calculations in Fig. 6(a) for the coherent intensity, compared to the plane-wave reference solutions obtained by inserting (8) in (2). Figure 7(b) depicts the absolute error of the noncoherent intensity in the far-field zone normalized to r . Here, the GBS result has been calculated by propagating the beams to the far zone, while the reference solution has been calculated directly from the plane-wave solution via (12b).

VI. CONCLUDING REMARKS

We presented a new scheme for rough surface scattering that is based on the Gaussian beam summation (GBS) method. The scheme localizes the scattering process in terms of stochastic GB2GB scattering matrices and deterministic GB propagators. It benefits from the simplicity and accuracy of the latter, and can therefore be used in various applications involving propagation in complex environments and in inverse imaging.

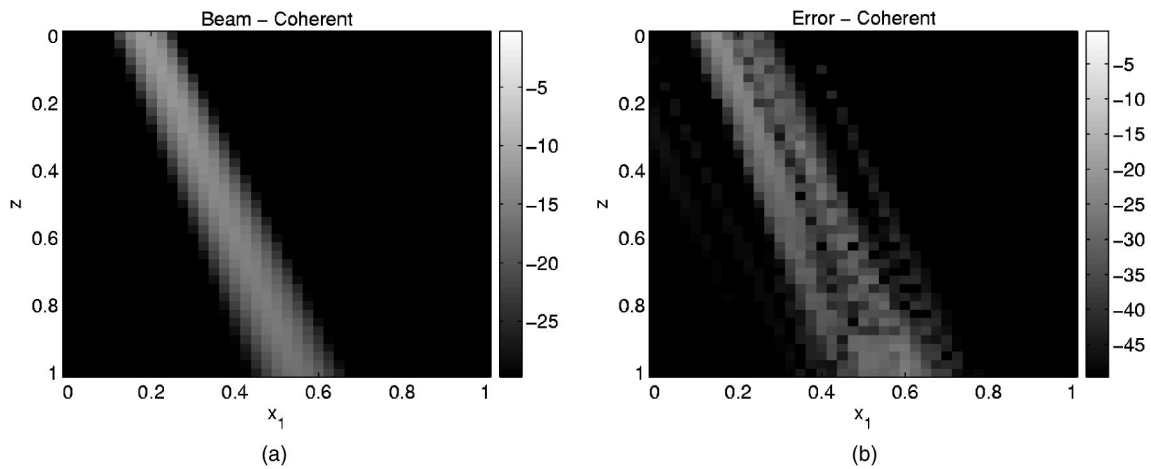


FIG. 6. Beam summation representation of the coherent intensities for the problem in Figs. 4 and 5, for the (x_1, z) plane for $x_2=0$. (a) Coherent intensity in dB calculated using the GBS. (b) Absolute error in dB compared to plane-wave reference solution.

The GB2GB scattering matrices in (51), are the beam equivalent of the known plane-wave scattering strength. They can be obtained either analytically, by applying the frame decomposition in conjunction with existing rough surface theories, as was done in the analytic example in Sec. V, or by applying it to measured (McDaniel, 1993) or numerically computed (Warnick and Chew, 2001) scattering moments. A more direct way to ascertain these matrices is to directly measure them, by way of inverse imaging.

In this work we presented the general formulation and applied it to a weakly rough surface, for which the scattering moments are known in closed form via the small perturbation approximation. The accuracy of the GBS representation for the calculation of the propagated coherent and noncoherent intensities has been demonstrated via an example in a homogeneous domain. In more complex scenarios with inhomogeneous media, the beam propagators provide a more efficient and accurate alternative to their plane-wave counterparts (Heilpern, 2004).

From a wave theoretic point of view, the GB2GB matrices represent the phase-space footprint of the rough surface scattering process. In more general propagation scenarios,

e.g., the doubly rough surface waveguide with multiple reflection phenomena which is studied in (Gordon *et al.*, 2005), the phase-space representation combines the footprints of the stochastic multiple scattering events and the deterministic propagation between the surfaces.

ACKNOWLEDGMENTS

This paper is dedicated to the memory of our coauthor, the late Professor Reuven Mazar. E. Heyman acknowledges partial support by the Israel Science Foundation under Grant No. 216/02 and by the US–Israel Binational Science Foundation, Jerusalem, Israel, under Grant No. 9900448.

APPENDIX: ANALYTICAL EVALUATION OF 2D GAUSSIAN-TYPE INTEGRALS

Throughout the development of the analytical example, a common form of an integral is repeated. The integral is of the form of a 2D slowly varying function $f(\mathbf{k}_x)$ multiplied by a 2D Gaussian exponent. The general form is

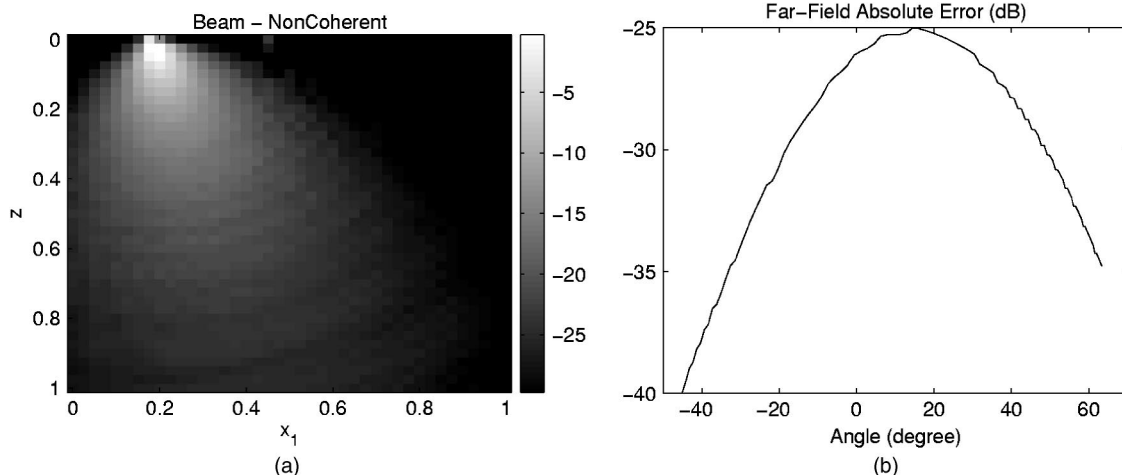


FIG. 7. Beam summation representation of the noncoherent intensities for the problem in Figs. 4 and 5, for the (x_1, z) plane for $x_2=0$. (a) Noncoherent intensity in dB calculated using the GBS. (b) Far-field normalized absolute error in dB compared to plane-wave reference solution, at $r=10b^i$.

$$I(\mathbf{k}'_x, \mathbf{k}''_x) = \int f(\mathbf{k}_x) e^{i(\mathbf{k}_x - \mathbf{k}'_x)^T \mathbf{A}(\mathbf{k}_x - \mathbf{k}'_x)/2} \times e^{i(\mathbf{k}_x - \mathbf{k}''_x)^T \mathbf{B}(\mathbf{k}_x - \mathbf{k}''_x)/2} e^{i\mathbf{k}_x \cdot \mathbf{x}} d^2 k_x, \quad (\text{A1})$$

where \mathbf{A} and \mathbf{B} are 2×2 symmetric matrices with positive definite imaginary parts.

Such integrals are obtained, usually, by the projection of a Gaussian beam on another function (e.g., another Gaussian beam of different parameters). Doing the saddle-point evaluation, one equates the first derivative of the exponent to zero to get the stationary point

$$\mathbf{k}_s = (\mathbf{A} + \mathbf{B})^{-1} (\mathbf{A}\mathbf{k}'_x + \mathbf{B}\mathbf{k}''_x - \mathbf{x}). \quad (\text{A2})$$

Using this stationary point, one can get the proper evaluation

$$I(\mathbf{k}'_x, \mathbf{k}''_x) = f(\mathbf{k}_s) \frac{2\pi i}{\sqrt{\det(\mathbf{A} + \mathbf{B})}} \exp \left[\frac{i}{2} (\mathbf{k}'_x - \mathbf{k}''_x)^T \times (\mathbf{A} + \mathbf{B})^{-1} \mathbf{A} \mathbf{B} (\mathbf{k}'_x - \mathbf{k}''_x) - \frac{i}{2} \mathbf{x}^T (\mathbf{A} + \mathbf{B})^{-1} \mathbf{x} + i\mathbf{x}^T (\mathbf{A} + \mathbf{B})^{-1} (\mathbf{A}\mathbf{k}'_x + \mathbf{B}\mathbf{k}''_x) \right], \quad (\text{A3})$$

where we take the root $\sqrt{\det(\mathbf{A} + \mathbf{B})}$ which has a positive imaginary part.

- Babich, V. M., and Popov, M. M. (1989). "Gaussian summation method (review)," *Radiophys. Quantum Electron.* **39**, 1063–1081.
- Bastiaans, M. J. (1980). "The expansion of an optical signal into a discrete set of Gaussian beams," *Optik* **57**(1), 95–102.
- Beckmann, P., and Spizzichino, A. (1987). *The Scattering of Electromagnetic Waves from Rough Surfaces* (Artech House, Norwood, MA).
- Červený, V., Popov, M. M., and Psencik, I. (1982). "Computation of wave

- fields in inhomogeneous media—Gaussian beam approach," *Geophys. J. R. Astron. Soc.* **70**, 109–128.
- Ellis, D. D., Preston, J. R., and Urban, H. G. (eds). (1993). *Ocean Reverberation* (Kluwer, Dordrecht).
- Etter, P. C. (2001). "Recent advances in underwater acoustic modelling and simulation," *J. Sound Vib.* **240**(2), 351–383.
- Etter, P. C. (2003). *Underwater Acoustics Modeling and Simulations*, 3rd ed. (Spon, New York).
- Gordon, G., Heyman, E., and Mazar, R. (2005). "Phase-space beam summation analysis of rough surface waveguide," *J. Acoust. Soc. Am.*, **117**, 1922–1932.
- Heilpern, T. (2004). "Beam summation algorithm for radiation in stratified media," M. Phil. thesis, School of Electrical Engineering, Tel Aviv University, Israel.
- Lugara, D., Letrou, C., Shlivinski, A., Heyman, E., and Boag, A. (2003). "The frame based Gaussian beam summation method: Theory and Application," *Radio Sci.* **38**(2), 1–15.
- Mazar, R., and Felsen, L. B. (1989). "Stochastic geometrical theory of diffraction (SGTD)," *J. Acoust. Soc. Am.* **86**(6), 2292–2308.
- McDaniel, S. T. (1993). "Sea surface reverberation: A review," *J. Acoust. Soc. Am.* **94**(4), 1905–1922.
- Melamed, T. (1997). "Phase space beam summation: A local spectrum analysis of time dependent radiation," *J. Electromagn. Waves Appl.* **11**, 739–773.
- Porter, M. B. (1997). "Modeling sound propagation in the ocean," in *Computational Wave Propagation*, IMA Volumes in Mathematics and its Applications, edited by B. Engquist and G. Kriegsmann (Springer, New York), pp. 197–215.
- Porter, M. B., and Bucker, H. P. (1987). "Gaussian beam tracing for computing ocean acoustic fields," *J. Acoust. Soc. Am.* **82**, 1349–1359.
- Shlivinski, A., Heyman, E., Boag, A., and Letrou, C. (2004). "A phase-space beam summation formulation for wideband radiation," *IEEE Trans. Antennas Propag.* **52**, 2042–2056.
- Steinberg, B. Z., Heyman, E., and Felsen, L. B. (1991). "Phase-space beam summation for time-harmonic radiation from large apertures," *J. Opt. Soc. Am. A* **8**, 41–59.
- Voronovich, A. (1999). *Wave Scattering from Rough Surfaces* (Springer, Berlin).
- Warnick, K. F., and Chew, W. C. (2001). "Numerical simulation methods for rough surface scattering," *Waves Random Media* **11**(1), 1–30.

Phase-space beam summation analysis of rough surface waveguide

Goren Gordon and Ehud Heyman

Department of Physical Electronics, Tel Aviv University, Tel Aviv 69978, Israel

Reuven Mazar^{a)}

Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

(Received 30 August 2004; revised 14 December 2004; accepted 18 December 2004)

A Gaussian beam summation (GBS) formulation is introduced for a doubly rough boundary waveguide, wherein the coherent and incoherent scattered fields are decomposed into a discrete phase-space summation of Gaussian beams (GB) that emanate from the rough surfaces in all directions. The scheme involves deterministic GB propagators and stochastic GB-to-GB (GB2GB) scattering matrices for the coherent and incoherent fields, where each scattered beam is propagated inside the waveguide and is scattered again from the rough boundaries. The GB2GB matrices are calculated from the statistical moments of the scattering amplitude, which are given either analytically or empirically. An analytical and numerical example for a waveguide with weak boundary roughness is presented and discussed. The formulation reveals explicitly the phase-space footprint of the stochastic multiple scattering process at the rough boundaries, thus providing a cogent physical interpretation and an effective mathematical representation to the field. The formulation also accommodates the receiver's pattern in the same phase-space format. Bistatic reverberations inside a rough surface waveguide as a function of the range and of the source and the receiver directions are thus examined as an implementation example. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1858151]

PACS numbers: 43.30.Hw, 43.20.Fn [SLB]

Pages: 1922–1932

I. INTRODUCTION

Calculating the field inside a doubly rough surface waveguide arises naturally in the characterization of transport properties related to, for instance, sonar, shallow-water waves, geophysical probing, remote sensing, radiowave propagation, as well as optical waveguides and fibers (Sheng, 1990; Etter, 2001, 2003, and references therein). Calculating the guided field may be performed via many approaches such as plane-wave spectrum (Voronovich, 1999), stochastic modes (Snchez-Gil *et al.*, 1998), rays, and other methods (Etter, 2001, 2003). However, keeping track of the field through the nonspecular, noncoherent multiple reflections at the waveguide rough boundaries, and treating these reflected wave species collectively to form the guided field, poses a problem of considerable complexity, which increases even further if the propagation medium is inhomogeneous.

For the coherent field, the plane-wave formulation is based on the mean reflection coefficient in the specular direction. The resulting series of multiple interactions may therefore be treated collectively, leading to a closed-form modal solution (Voronovich, 1999). Such collective procedure is not applicable for the noncoherent field due to the nonspecular nature of the interactions. The latter problem may be addressed via perturbations, as was done in Schmidt and Kuperman (1995) for low frequencies in stratified waveguides with rough surface boundaries. The modal solu-

tion may be formulated, alternatively, by using a coupled-modes approach, but for high frequencies and inhomogeneous media, the coupling matrices become large and unpractical (Beenakker, 1997). Alternatively, the noncoherent field can be described by tracking the series of multiple nonspecular, noncoherent reflections. The complexity of this approach increases with the interaction order, and thereby with the range, but it may be reduced by applying localization using ray representations (Mazar and Felsen, 1989). The latter may also be extended to range-dependent nonuniform waveguides, but, on the other hand, since they are stripped of the spectral flash, they are sensitive to nonphysical artifacts such as caustics and shadow zones.

To overcome the difficulties of the ray and spectral approaches, we utilize here the Gaussian beam summation (GBS) representation of the rough surface scattering, introduced by us in Gordon *et al.* (2005). This formulation is structured on the localized ray skeleton but retains the spectral characteristics of the plane-wave formulations. As a result, it combines the localization and algorithmical properties of the ray representation with the uniform features of the plane-wave representation.

The GBS method was introduced originally for tracking source-excited wave fields in deterministic configurations (see the review articles in Babich and Popov, 1989; Nowack, 2003; Heyman and Felsen, 2001). Its extension in Gordon *et al.* (2005) to rough surface scattering applications is managed by the coherent and noncoherent GB-to-GB (GB2GB) scattering matrices, where the incident field is decomposed into GBs and each GB is then scattered coherently and non-

^{a)}Deceased.

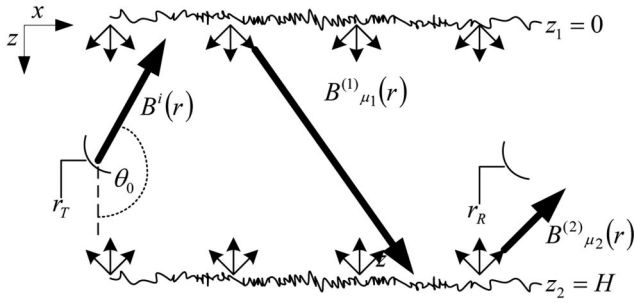


FIG. 1. The Gaussian beam summation (GBS) in a waveguide with rough surface boundaries at $z_{1,2}$ and roughness $h_{1,2}$. The coordinate frame is $\mathbf{r} = (x_1, x_2, z) = (\mathbf{x}, z)$ with “downward” coordinate z . The transmitter T and receiver R are located at $\mathbf{r}_T = (x_T, z_T)$ and $\mathbf{r}_R = (x_R, z_R)$, respectively. In the examples it is assumed that both have a Gaussian pattern and are looking toward the z_1 surface as depicted. The scattered fields from $z_{1,2}$ are decomposed into discrete sets of GB propagators $B_{\mu_{1,2}}^{(1,2)}$, where $\mu_{1,2}$ are the phase-space indices. The rough surface scattering is described by the GB2GB matrices $\mathbf{S}_c^{(1,2)}$ and $\mathbf{S}_{nc}^{(1,2)}$ for the coherent and the noncoherent fields, respectively.

coherently into a discrete spectrum of GBs. The formulation is based on the frame-based phase-space GBS representation in Shlivinski *et al.* (2004), where the scattered field is expressed as a sum of beams emanating from a phase-space lattice of points and directions on the rough surface, as depicted in Fig. 1 and summarized briefly in Sec. II.

In this paper the GBS approach of Gordon *et al.* (2005) for rough surface scattering is extended to accommodate propagation in inhomogeneous waveguides with rough surface boundaries, which are characterized by multiple interactions of the coherent and incoherent fields with the rough surfaces (see Figs. 1 and 5). The formulation utilizes GBs to propagate these fields locally within the medium, and the GB2GB scattering matrix to acquire the relevant GBs scattered from the rough surfaces. An important feature of this new representation is that it reveals explicitly the local phase-space footprint of the stochastic scattering process at the rough boundaries, thus providing a cogent physical interpretation to the field propagation within the waveguide. These localization features are utilized in the algorithm to obtain an effective mathematical representation of the problem, where we have also neglected multiple noncoherent scattering.

Concerning the presentation layout, we start in Sec. II with a brief overview of Gordon *et al.* (2005) for the GBS analysis of rough surface scattering. The application of the GBS method to the doubly rough surface waveguide problem is presented in Sec. III. Two alternative representations are presented: a series of successive interactions and a collective solution (Secs. III B and III C, respectively). An important feature of the formulation is the phase-space localization implied by the collimated propagators and by the GB2GB scattering matrices. The implication of the localization on the GBS algorithm, as well as the “proper” choice of the algorithm parameters, are discussed in Secs. III D–III F. Finally, all these properties are explored via a comprehensive numerical example in Sec. IV.

The theory of Sec. III is phrased in terms of field observables. In Sec. V it is rephrased in terms of the coupling

into the receiving antenna which is characterized by a receiving pattern. As an example the theory is applied in Sec. V C for the calculations of the bistatic reverberations in a two-antenna system. Finally, the extension of the method to an inhomogeneous waveguiding configuration is briefly discussed in Sec. VI, where we mainly consider the analysis of possible surface ducts near the rough surface boundaries. The presentation ends with concluding comments in Sec. VII.

II. PHASE-SPACE GAUSSIAN BEAM SUMMATION ANALYSIS OF ROUGH SURFACE SCATTERING: A BRIEF REVIEW

We consider the scattered wave field $u(\mathbf{r})$ in the half-space $z > h(\mathbf{x})$ due to a rough surface boundary $z = h(\mathbf{x})$, where z is the vertical coordinate in a 3D coordinate frame $\mathbf{r} = (\mathbf{x}, z)$ and $\mathbf{x} = (x_1, x_2)$ is the horizontal coordinate. The scattering observables are the average field and/or the average intensity, where the averaging is performed over the ensemble of realizations. The average field defines the so-called coherent field. The average intensity consists of two terms, the coherent and the noncoherent intensities, where the latter describes the intensity fluctuations due to rough surface. Our goal is to calculate and characterize these observables via the GBS. This will be done by calculating the coherent and the noncoherent fields u_c and u_{nc} , respectively, which may be understood as mathematical realizations that give rise to the statistical observables noted above (e.g., $|u_{nc}|^2$ describes the noncoherent intensity).

A. The phase-space Gaussian beam summation method

We start with a brief overview of this method and then apply it in Sec. II B to the rough surface problem. The method provides a rigorous framework that decomposes any field $u_0(\mathbf{x})$ in an aperture of a real or virtual source to a set of GBs that emanates from a discrete set of points in the aperture into a discrete set of directions. The starting point is the expansion of the aperture field $u_0(\mathbf{x})$ in a plane of constant z (say $z = 0$) using shifted and modulated window functions of the form

$$\psi_{\boldsymbol{\mu}}(\mathbf{x}) = \psi(\mathbf{x} - \mathbf{x}_m) e^{i\mathbf{k}_{x\mathbf{n}} \cdot (\mathbf{x} - \mathbf{x}_m)}, \quad \mathbf{x}_m = \mathbf{m}\bar{x}, \quad \mathbf{k}_{x\mathbf{n}} = \mathbf{n}\bar{k}_x, \quad (1)$$

where $\boldsymbol{\mu} = (\mathbf{m}, \mathbf{n}) = (m_1, m_2, n_1, n_2)$ is an index and ψ is a proper window function. The $\boldsymbol{\mu}$ elements are centered about the phase-space lattice points $(\mathbf{x}_m, \mathbf{k}_{x\mathbf{n}})$. This set of functions needs to be complete or even overcomplete to be able to expand any aperture function. When it is overcomplete it is called a windowed Fourier transform (WFT) frame. A necessary condition is that the lattice is overcomplete such that the spatial and spectral unit cell dimensions (\bar{x}, \bar{k}_x) satisfy

$$\bar{x}\bar{k}_x = 2\pi\nu, \quad \text{where } \nu < 1. \quad (2)$$

ν is the frame overcompleteness (or oversampling) factor. The frame is overcomplete for $\nu < 1$ and it is complete in the Gabor limit $\nu \uparrow 1$. The overcomplete frame in (1) can be used to expand the data $u_0(\mathbf{x})$ as

$$u_0(\mathbf{x}) = \sum_{\mu} a_{\mu} \psi_{\mu}(\mathbf{x}). \quad (3)$$

Since the frame is overcomplete, the coefficient set a is nonunique, i.e., there are many sets that can be used in (3) to reconstruct $u_0(\mathbf{x})$. A desirable set is obtained by projecting the data onto the dual frame $\varphi_{\mu}(\mathbf{x})$ via

$$a_{\mu} = \langle u_0(\mathbf{x}), \varphi_{\mu}(\mathbf{x}) \rangle = \int d^2x u_0(\mathbf{x}) \varphi_{\mu}^*(\mathbf{x}), \quad (4)$$

where $\varphi_{\mu}(\mathbf{x})$ has the same form as the frame ψ_{μ} in (1) except that the window $\psi(\mathbf{x})$ is now replaced by its dual $\varphi(\mathbf{x})$. Equation (4) can readily be identified as a WFT of u_0 with respect to the window φ , hence the name WFT frames. In view of (3) and (4), ψ and φ are also termed the synthesis and analysis windows, respectively. In general, $\varphi(\mathbf{x})$ needs to be calculated for a given $\psi(\mathbf{x})$ and (\bar{x}, \bar{k}_x) , but for small ν it can be found analytically [see (6)].

Henceforth, we use a matched Gaussian of the form

$$\psi(\mathbf{x}) = e^{-k|\mathbf{x}|^2/2b}, \quad b = k\bar{x}/\bar{k}_x, \quad (5)$$

where k is the wave number in the ambient medium and $b > 0$ is a parameter. The condition on b in (5) defines the window width $W_0 = \sqrt{b/k}$ such that ψ is matched to the lattice as discussed in [Gordon *et al.*, 2005, Eq. (19)]. In wave propagation applications we typically choose b such that $kb \gg 1$ so that W_0 is wide on the wavelength scale, giving rise to collimated GB propagators [see (7)]. Finally, one should note that choosing ν and b determines the lattice dimensions (\bar{x}, \bar{k}_x) via (1) and (5).

The overcompleteness parameter ν is chosen as a trade-off between analytic simplicity and numerical efficiency. The latter implies that ν should be as large as possible, while the former implies that it should be sufficiently small to yield a convenient dual window. Indeed, for $\nu < 0.5$, φ that corresponds to the matched Gaussian windows (5) is given approximately by Gordon *et al.* (2005), Eq. (32)

$$\varphi(\mathbf{x}) \approx \nu^2 (k/\pi b) e^{-k|\mathbf{x}|^2/2b}. \quad (6)$$

For $\nu=0.5$ the L_2 error in (6) is $\sim 8\%$, while for $\nu=0.3$ it is $\sim 1\%$, (Shlivinski *et al.*, 2004, Fig. 4). Thus, $\nu=0.5$ provides a reasonable choice for the present application.

Equation (3) may be propagated now away from the aperture, giving

$$u(\mathbf{r}) = \sum_{\mu} a_{\mu} B_{\mu}(\mathbf{r}), \quad (7)$$

where each $B_{\mu}(\mathbf{r})$ is the propagated field due to the aperture distribution $\psi_{\mu}(\mathbf{x})$. For the windows in (5) these propagators are GBs whose properties are discussed in Gordon *et al.*, 2005, Eqs. (32)–(40). They emerge from the points \mathbf{x}_m in the aperture plane in directions that are determined by $\mathbf{k}_{x,m}$ such that their angles with respect to the x_j axes, $j=1, 2$, are given by $\cos^{-1}(k_{x,m}/k)$. For $|\mathbf{k}_{x,m}| > k$, the fields $B_{\mu}(\mathbf{r})$ decay away from the aperture and hence are not included in the summation (7).

B. Rough surface scattering using GBS: The GB2GB scattering matrix

Applying the GBS to the rough surface scattering discussed above, the coherent and the noncoherent fields are expanded as (Gordon *et al.*, 2005)

$$u_c^s(\mathbf{r}) = \sum_{\mu} S_c(u^i, \psi_{\mu}) B_{\mu}(\mathbf{r}), \quad (8a)$$

$$u_{nc}^s(\mathbf{r}) = \sum_{\mu} S_{nc}(u^i, \psi_{\mu}) B_{\mu}(\mathbf{r}), \quad (8b)$$

where $S_c(u^i, \psi_{\mu})$ and $S_{nc}(u^i, \psi_{\mu})$ are the coherent and noncoherent scattering matrices from the incident field u^i to the scattering propagator $B_{\mu}(\mathbf{r})$. For a given surface they can be calculated via

$$S_c(u^i, \psi_{\mu}) = \frac{1}{(2\pi)^2} \int \tilde{u}^i(\mathbf{k}_x) \bar{V}(\mathbf{k}_x) \tilde{\varphi}_{\mu}^*(\mathbf{k}_x) d^2k_x, \quad (9a)$$

$$S_{nc}(u^i, \psi_{\mu}) = \frac{1}{(2\pi)^2} \int e^{-i\mathbf{k}_x \cdot \mathbf{x}^i} \times \sqrt{\int \sigma(\mathbf{k}_x, \mathbf{k}'_x) |\tilde{u}^i(\mathbf{k}'_x)|^2 d^2k'_x} \tilde{\varphi}_{\mu}^*(\mathbf{k}_x) d^2k_x, \quad (9b)$$

where $\tilde{u}^i(\mathbf{k}_x)$ is the spectral counterpart of $u^i(\mathbf{x})$, \mathbf{x}^i is the center of mass of $u^i(\mathbf{x})$, and $\tilde{\varphi}_{\mu}(\mathbf{k}_x)$ is the spectral counterpart of φ_{μ} . Here, $\bar{V}(\mathbf{k}_x)$ and $\sigma(\mathbf{k}_x, \mathbf{k}'_x)$ are the mean spectral reflection coefficient and the spectral scattering cross section. For certain classes of surfaces, these scattering characteristics of $h(\mathbf{x})$ can be found analytically; otherwise, they can be found empirically from the measured observables as discussed in Gordon *et al.*, 2005, Sec. V.

Finally, since u^i can be decomposed also into incident GBs, the scattered field can be expressed in terms of GB to GB (GB2GB) scattering matrices, which are given as in (9) except that u^i is now replaced by an incident beam B^i .

III. PHASE-SPACE ANALYSIS OF ROUGH SURFACE WAVEGUIDE

The GB2GB scattering representation derived above provides the theoretical framework for the analysis of rough surface waveguides where the field undergoes a multiple scattering process between the rough surface walls. In this formulation, the field that is scattered by the wall in a given iteration is decomposed into a set of GBs that, in turns, constitutes the incident field on the other wall at the next iteration. This solution procedure can be addressed either iteratively, via an interactions series approach, or collectively (see Secs. III B and III C, respectively). Conceptually, these solution approaches are similar to plane-wave procedures in Voronovich (1999) [Chap. 10]. Yet, the beam formulation is not only *a priori* discrete, but it may also accommodate range-dependent inhomogeneous media.

A. The phase-space beam setup

The problem setup in Fig. 1 consists of a homogeneous waveguide, with two rough surface boundaries $z=z_1+h_1(\mathbf{x})$ and $z=z_2+h_2(\mathbf{x})$, where $z_1 < z_2$ while $h_{1,2}$ denote the surface roughnesses which are assumed to be statistically uncorrelated and \mathbf{x} -independent. As before, z is the vertical depth coordinate in a 3D coordinate frame $\mathbf{r}=(\mathbf{x},z)$ and $\mathbf{x}=(x_1,x_2)$ is the horizontal (or range) coordinate. Note that in principle the GBS formulation below may be applied to waveguides with weak range dependence of the propagation medium and of the surface roughness.

Since the two rough surfaces are noncorrelated, each one may be treated independently from the other. We thus construct two overcomplete lattices of Gaussian beams on the two rough surface boundaries, tagged by the indices $\boldsymbol{\mu}_j$ where $j=1, 2$ for z_1 and z_2 , respectively (Fig. 1). Since the frames on the z_1 and z_2 boundaries are not necessarily the same, we denote them as $(\bar{x}^{(j)}, \bar{k}_x^{(j)})$ and $\psi_{\boldsymbol{\mu}_j}^{(j)}(\mathbf{x})$, and the corresponding GB propagators as $B_{\boldsymbol{\mu}_j}^{(j)}(\mathbf{r})$. The field impinging on the j th boundary is thus expressed as a summation of

GBs $B_{\boldsymbol{\mu}_{j'}}^{(j')}(\mathbf{r})$ arriving from the other j' boundary, and the scattered field is then described in terms of the GB2GB scattering matrices $S^{(j'j)}$ from $B_{\boldsymbol{\mu}_{j'}}^{(j')}(\mathbf{r})$ to $B_{\boldsymbol{\mu}_j}^{(j)}(\mathbf{r})$. The field inside the waveguide is then expressed by summing up all the multiply scattered beams.

B. Interaction series approach

The summation process starts by expressing the source excited field using GBs $u^i(\mathbf{r}) = \sum_{\boldsymbol{\mu}_0} a_{\boldsymbol{\mu}_0} B_{\boldsymbol{\mu}_0}^{(0)}(\mathbf{r})$ tagged by index $\boldsymbol{\mu}_0$ [the beam set $B_{\boldsymbol{\mu}_0}^{(0)}$ which is used to expand the source field is in general different than $B_{\boldsymbol{\mu}_j}^{(j)}$; see the discussion in Gordon *et al.*, 2005, Eq. (41), and Fig. 5(a)]. Without loss of generality, we assume that the source radiates only toward z_1 , where each $B_{\boldsymbol{\mu}_0}^{(0)}$ excites a spectrum of scattered beams $B_{\boldsymbol{\mu}_1}^{(1)}(\mathbf{r})$ residing on the lattice $\boldsymbol{\mu}_1$ that propagates toward z_2 . At z_2 each $B_{\boldsymbol{\mu}_1}^{(1)}$ scatter into a spectrum of beams $B_{\boldsymbol{\mu}_2}^{(2)}(\mathbf{r})$ on the lattice $\boldsymbol{\mu}_2$, and vice versa. This process is formalized in the following scheme:

$$u(\mathbf{r}) = u^i(\mathbf{r}) + \sum_{\boldsymbol{\mu}_0} \sum_{\boldsymbol{\mu}'_1} a_{\boldsymbol{\mu}_0} S(B_{\boldsymbol{\mu}_0}^{(0)}|_{z_1}, \psi_{\boldsymbol{\mu}'_1}^{(1)}) B_{\boldsymbol{\mu}'_1}^{(1)}(\mathbf{r}) + \sum_{\boldsymbol{\mu}_0} \sum_{\boldsymbol{\mu}'_1} \sum_{\boldsymbol{\mu}'_2} a_{\boldsymbol{\mu}_0} S(B_{\boldsymbol{\mu}_0}^{(0)}|_{z_1}, \psi_{\boldsymbol{\mu}'_1}^{(1)}) S(B_{\boldsymbol{\mu}'_1}^{(1)}|_{z_2}, \psi_{\boldsymbol{\mu}'_2}^{(2)}) B_{\boldsymbol{\mu}'_2}^{(2)}(\mathbf{r}) \\ + \sum_{\boldsymbol{\mu}_0} \sum_{\boldsymbol{\mu}'_1} \sum_{\boldsymbol{\mu}'_2} \sum_{\boldsymbol{\mu}''_1} a_{\boldsymbol{\mu}_0} S(B_{\boldsymbol{\mu}_0}^{(0)}|_{z_1}, \psi_{\boldsymbol{\mu}'_1}^{(1)}) S(B_{\boldsymbol{\mu}'_1}^{(1)}|_{z_2}, \psi_{\boldsymbol{\mu}'_2}^{(2)}) S(B_{\boldsymbol{\mu}'_2}^{(2)}|_{z_1}, \psi_{\boldsymbol{\mu}''_1}^{(1)}) B_{\boldsymbol{\mu}''_1}^{(1)}(\mathbf{r}) + \dots, \quad (10)$$

where

- (i) The first term is the incident field that propagates toward z_1 .
- (ii) The second term represents the scattering of the incident beams $B_{\boldsymbol{\mu}_0}^{(0)}$ evaluated at $z=z_1$ toward the set of beams $B_{\boldsymbol{\mu}'_1}^{(1)}$ on the lattice $\boldsymbol{\mu}_1$ with $S(B_{\boldsymbol{\mu}_0}^{(0)}|_{z_1}, \psi_{\boldsymbol{\mu}'_1}^{(1)})$ denoting the GB2GB scattering matrix (the prime in $\boldsymbol{\mu}'_1$ tags the first scattering at z_1 ; higher-order scattering are tagged by double primes, etc.). Note that in this term and in the following ones the sum $\sum_{\boldsymbol{\mu}_0} a_{\boldsymbol{\mu}_0} S(B_{\boldsymbol{\mu}_0}^{(0)}|_{z_1}, \psi_{\boldsymbol{\mu}'_1}^{(1)})$ can be replaced by $S(u^i|_{z_1}, \psi_{\boldsymbol{\mu}'_1}^{(1)})$, which represents the scattering matrix from u^i evaluated at $z=z_1$ toward the set of beams $B_{\boldsymbol{\mu}'_1}^{(1)}$. We use, however, the series format since it is expressed explicitly in terms of incident GBs and the GB2GB scattering matrix, whereas u^i is in general not Gaussian.
- (iii) The third term represents the scattering of each of the beams $B_{\boldsymbol{\mu}'_1}^{(1)}$ evaluated at z_2 toward the set of beams $B_{\boldsymbol{\mu}'_2}^{(2)}$ on the lattice $\boldsymbol{\mu}_2$, with $S(B_{\boldsymbol{\mu}'_1}^{(1)}|_{z_2}, \psi_{\boldsymbol{\mu}'_2}^{(2)})$ representing the GB2GB scattering matrix of boundary z_2 .

The iterative process in Eq. (10) can be expressed, con-

veniently, in a matrix format. We define the ‘‘propagators vectors’’ $\mathbf{B}^{(j)}(\mathbf{r}) = B_{\boldsymbol{\mu}_j}^{(j)}(\mathbf{r})$, the ‘‘waveguide interaction matrices’’ $\mathbf{S}^{(jj')}$, and the ‘‘source scattering row vector’’ $\mathbf{S}^{(0)}$

$$\mathbf{B}^{(j)}(\mathbf{r}) = B_{\boldsymbol{\mu}_j}^{(j)}(\mathbf{r}), \quad j=1,2 \quad (11)$$

$$\mathbf{S}^{(jj')} = S(B_{\boldsymbol{\mu}_j}^{(j)}|_{z_j}, \psi_{\boldsymbol{\mu}_{j'}}^{(j')}), \quad (j,j') = (1,2) \text{ or } (2,1) \quad (12)$$

$$\mathbf{S}^{(0)} = S(u^i, \psi_{\boldsymbol{\mu}_1}^{(1)}) = \sum_{\boldsymbol{\mu}_0} a_{\boldsymbol{\mu}_0} S(B_{\boldsymbol{\mu}_0}^{(0)}|_{z_1}, \psi_{\boldsymbol{\mu}_1}^{(1)}). \quad (13)$$

The vectors $\mathbf{B}^{(j)}$ consist of all phase-space beams emerging from the j boundary. The matrices $\mathbf{S}^{(jj')}$ account for the propagation from boundary j to boundary j' and for the subsequent scattering there. These matrices are source independent and can be calculated *a priori* for a given waveguide configuration (uniform as well as nonuniform). Finally, in the definition of $\mathbf{S}^{(0)}$ it has been assumed that the source illuminates surface 1 only.

The iterative solution in (10) may be rewritten now in a matrix format

$$u(\mathbf{r}) = u^i(\mathbf{r}) + \mathbf{S}^{(0)} \mathbf{B}^{(1)}(\mathbf{r}) + \mathbf{S}^{(0)} \mathbf{S}^{(12)} \mathbf{B}^{(2)}(\mathbf{r}) \\ + \mathbf{S}^{(0)} \mathbf{S}^{(12)} \mathbf{S}^{(21)} \mathbf{B}^{(1)}(\mathbf{r}) + \dots \quad (14)$$

Note that the matrices involve all the lattice points on the boundaries and are therefore infinite. Yet, due to beam local-

ization, the $\mathbf{S}^{(jj')}$ are rather sparse (each $\boldsymbol{\mu}_j$ incident beam gives rise to a small subset of $\boldsymbol{\mu}_{j'}$ scattered beams), and therefore they, as well as the iterative series in (14), may be truncated depending on the propagation range (see the discussion in Sec. III E).

C. A collective representation

The interactions series in (10) can be summed up collectively. Noting that (10) [or (14)] consists of two beam sets: $\mathbf{B}^{(1)}(\mathbf{r})$ that propagate from z_1 to z_2 and $\mathbf{B}^{(2)}(\mathbf{r})$ that propagate from z_2 to z_1 , these expressions can be rewritten as

$$u(\mathbf{r}) = u^i(\mathbf{r}) + \sum_{\boldsymbol{\mu}_1} F_{\boldsymbol{\mu}_1}^{(1)} B_{\boldsymbol{\mu}_1}^{(1)}(\mathbf{r}) + \sum_{\boldsymbol{\mu}_2} F_{\boldsymbol{\mu}_2}^{(2)} B_{\boldsymbol{\mu}_2}^{(2)}(\mathbf{r}) \quad (15a)$$

$$= u^i(\mathbf{r}) + \mathbf{F}^{(1)} \mathbf{B}^{(1)}(\mathbf{r}) + \mathbf{F}^{(2)} \mathbf{B}^{(2)}(\mathbf{r}). \quad (15b)$$

One may readily show that the coefficients vectors $\mathbf{F}^{(j)} = F_{\boldsymbol{\mu}_j}^{(j)}$ satisfy the equation set

$$\mathbf{F}^{(1)} = \mathbf{S}^{(0)} + \mathbf{F}^{(2)} \mathbf{S}^{(21)}, \quad (16a)$$

$$\mathbf{F}^{(2)} = \mathbf{F}^{(1)} \mathbf{S}^{(12)}. \quad (16b)$$

Equation (16) can be solved via iterations

$$\mathbf{F}^{(1)} = \mathbf{S}^{(0)} + \mathbf{S}^{(0)} \mathbf{S}^{(12)} \mathbf{S}^{(21)} + \dots, \quad (17a)$$

$$\mathbf{F}^{(2)} = \mathbf{S}^{(0)} \mathbf{S}^{(12)} + \mathbf{S}^{(0)} \mathbf{S}^{(12)} \mathbf{S}^{(21)} \mathbf{S}^{(12)} + \dots, \quad (17b)$$

thus recovering the field solution in (10).

Alternatively, a formal ‘‘collective solution’’ of (16) is given by

$$\mathbf{F}^{(1)} = \mathbf{S}^{(0)} [\mathbf{I} - \mathbf{S}^{(12)} \mathbf{S}^{(21)}]^{-1}, \quad (18a)$$

$$\mathbf{F}^{(2)} = \mathbf{S}^{(0)} [\mathbf{I} - \mathbf{S}^{(12)} \mathbf{S}^{(21)}]^{-1} \mathbf{S}^{(12)}, \quad (18b)$$

where \mathbf{I} is the identity matrix, and finally

$$u(\mathbf{r}) = u^i(\mathbf{r}) + \mathbf{S}^{(0)} [\mathbf{I} - \mathbf{S}^{(12)} \mathbf{S}^{(21)}]^{-1} \times \{\mathbf{B}^{(1)}(\mathbf{r}) + \mathbf{S}^{(12)} \mathbf{B}^{(2)}(\mathbf{r})\}. \quad (19)$$

As noted above, the $\mathbf{S}^{(jj')}$ matrices are in principle infinite but, depending on the propagation range, they may be truncated so that the inverse matrices in (18) can be calculated. Although the matrices’ dimension grows with the range, one may still utilize the matrices’ sparseness in order to reduce the complexity, as will be discussed in Sec. III E.

D. Coherent and noncoherent fields

As discussed in Sec. II, the total field is the sum of the coherent and noncoherent fields. Thus, in the formulations in Secs. III B and III C, we may express the scattering matrices as a sum of the coherent and noncoherent matrices $\mathbf{S} = \mathbf{S}_c + \mathbf{S}_{nc}$. However, depending on the observables, one wishes to distinguish these two contributions.

The coherent field is generated by coherent multiple scattering processes. It is therefore described by using only the coherent matrices \mathbf{S}_c in the interactions series (14) or in the collective solution (19).

For the noncoherent field (i.e., the average intensity as observable) one must take all the other combinations of co-

herent and noncoherent interactions, i.e., at least one \mathbf{S} matrix in each one of the additive terms in (14) must be \mathbf{S}_{nc} . A simple way to address this combinatorial problem is to set $\mathbf{S} = \mathbf{S}_c + \mathbf{S}_{nc}$ in (14) or (19) and then subtract the coherent field from the final result.

It should be noted, though, that \mathbf{S}_c contains mainly specular beams and thus only a few non-negligible elements, while \mathbf{S}_{nc} is spectrally wide and consists of nonspecular beam interactions and therefore of a large number of non-negligible elements (see Gordon *et al.*, 2004b, Fig. 5) and also Figs. 2 and 3 below). Consequently, the number of elements in the calculation of the noncoherent field via the interactions series (14) grows rapidly with the interaction order. However, in many cases the elements of \mathbf{S}_{nc} are weaker relative to those of \mathbf{S}_c , so that the n th additive term in the series can be described, to leading order, by *only one noncoherent interaction* where all the other $(n-1)$ interactions are coherent. The n th term is thus described as a sum of n permutations (e.g., for $n=3$: $\mathbf{S}_{nc}^{(0)} \mathbf{S}_c^{(12)} \mathbf{S}_c^{(21)} + \mathbf{S}_c^{(0)} \mathbf{S}_{nc}^{(12)} \mathbf{S}_c^{(21)} + \mathbf{S}_c^{(0)} \mathbf{S}_c^{(12)} \mathbf{S}_{nc}^{(21)}$ or, in short, $nc-c-c$, $c-nc-c$, and $c-c-nc$, whereas terms with two or more noncoherent interactions such as $nc-nc-c$, etc., are neglected). This reduces the total number of elements in the calculation of the noncoherent field. Each of these permutation yields a different physical structure, as will be demonstrated in the numerical example of Fig. 3.

E. Beam localization and computational aspects

In both the interactions series and collective approaches, the computational burden is divided into two aspects, one being the calculation of the propagators and then of the waveguide interaction matrices $\mathbf{S}^{(jj')}$, and the other is the summation of the beam contributions. Both can be considerably simplified by considering the beam localization.

The localization is affected via both the *propagation* and the *scattering* mechanisms. The latter implies that the scattering matrices are localized as shown in Gordon *et al.*, 2005, Fig. 5, where we explored the coherent and noncoherent scattering matrices for a given incident beam. The $\mathbf{S}_c^{(jj')}$ matrices represent essentially specular returns and are therefore localized about the phase-space footprint of the incident beams. For example, in homogeneous medium $\mathbf{S}_c^{(jj')}$ is non-negligible if $\boldsymbol{\mu}_j$ and $\boldsymbol{\mu}_{j'}$ are related via $\mathbf{k}_{x\mathbf{n}_j} \approx \mathbf{k}_{x\mathbf{n}_{j'}}$ and $\mathbf{x}_{\mathbf{m}_j} \approx \mathbf{x}_{\mathbf{m}_{j'}} + H \mathbf{k}_{x\mathbf{n}_j} k_{z\mathbf{n}_j}^{-1}$, where $H = z_2 - z_1$ and $k_{z\mathbf{n}_j} = \sqrt{k^2 - |\mathbf{k}_{x\mathbf{n}_j}|^2}$.

The $\mathbf{S}_{nc}^{(jj')}$ matrices are localized only about the *spatial* footprint of the incident beam but they have a wide spectral spread, which depends among other things on the correlation length [see the explicit example in Gordon *et al.* (2005)]. In homogeneous medium, they have non-negligible elements ‘‘for all’’ \mathbf{n}_j and $\mathbf{n}_{j'}$ while \mathbf{m}_j and $\mathbf{m}_{j'}$ are related via $\mathbf{x}_{\mathbf{m}_j} \approx \mathbf{x}_{\mathbf{m}_{j'}} + H \mathbf{k}_{x\mathbf{n}_j}^{-1} \mathbf{k}_{z\mathbf{n}_j}^{-1}$.

The sparseness of the $\mathbf{S}^{(jj')}$ matrices reduces the complexity and the overall computational cost of the calculation. Furthermore, the relevant matrix terms may be determined *a priori* by following the propagators’ trajectories, thus reducing the complexity further.

In the computational algorithm, we thus utilize both *source localization* and *observer localization* considerations. The former accounts for the phase-space footprint of the source by following the beams that are actually excited by the source through their interaction with the waveguide boundaries. The latter accounts for the phase-space footprint of the observer, i.e., the beams that pass near the observer and in the preferred direction of its receiving pattern.

To explain how the *source localization* is utilized, consider, as an example, the interaction series in (14). The incident field excites only a selected set of $\boldsymbol{\mu}_1$ in $\mathbf{S}^{(0)}$ which is known *a priori* for both $\mathbf{S}_c^{(0)}$ and $\mathbf{S}_{nc}^{(0)}$ in view of the discussion above. Each $\boldsymbol{\mu}_1$ beam then excites only a selected set of $\boldsymbol{\mu}_2$ in $\mathbf{S}^{(1)}$, and so forth. One notes that the *complexity* of the coherent field calculations increases slowly with the range due to the slow spectral spread of the coherent beam field. The complexity of the noncoherent field calculations increases faster, not only because of the permutations of $\mathbf{S}_{nc}^{(jj')}$ and $\mathbf{S}_c^{(jj')}$ that should be accounted for in each successive interaction (see the discussion in Sec. III D), but also since $\mathbf{S}_{nc}^{(jj')}$ is less sparse (spectrally wider) than $\mathbf{S}_c^{(jj')}$.

The *observer localization* is affected since the beam summation algorithm accounts only for the contribution of the beams that pass near the observation point (typically within a 3-beamwidth distance from that point). If, in addition, the observer has a receiving pattern, then the contributing beams should also be multiplied by this pattern, thus reducing further the number of relevant beams (see the discussion in Sec. V). For a given observation point and receiving pattern, these considerations define a submanifold in the phase space, denoted as the “geometrical observation manifold” (GOM) (Shlivinski *et al.*, 2004). The use of the GOM limits, *a priori*, the set of the phase-space propagators $B_{\boldsymbol{\mu}_1}^{(1)}(\mathbf{r})$ and $B_{\boldsymbol{\mu}_2}^{(2)}(\mathbf{r})$ that is needed at the observation point. The relevant propagator vectors $B_{\boldsymbol{\mu}_j}^{(j)}$ then define the relevant terms in the matrices that multiply propagator vectors $\mathbf{B}^{(j)}$ from the left in either the interaction series representations (14) or the collective representations (19), and so forth as we progress backward from the receiver to the source. One should recall here, again, the different localized scattering properties of $\mathbf{S}_c^{(jj')}$ and $\mathbf{S}_{nc}^{(jj')}$ as discussed above. These considerations will be demonstrated in the examples of Secs. IV and V C.

The matrix localization can be used to define the number of elements needed in the representation as a function of the observation range. Theoretically, the $\mathbf{S}^{(jj')}$ matrices are infinite since they involve all lattice points along the propagation range. However, one can take only a finite number of elements for a given range of calculation. Thus, the interaction series in (14) for the coherent field needs to be summed only until the specular reflections have passed the required range $|\mathbf{x}|$, whereas for the noncoherent field one should keep a few terms beyond this range due to the possible backward reflections associated with $\mathbf{S}_{nc}^{(jj')}$ (see the examples in Secs. IV and V C).

Similar considerations may also be used in the collective solution (19) where one can truncate the matrices at \mathbf{x}_m be-

yond the required range. This simplifies the calculation of the inverse matrix there.

F. Selecting the beam expansion parameters

The localization discussed in the preceding section is mainly affected by the properties of the frame propagators. By controlling these parameters one may change the expansion characteristics continuously from a ray-like solution to a plane-wave-like solution; hence, these parameters should be selected properly to match a given propagation configuration.

As discussed after (5) the only parameters that need to be chosen are the window parameters $b^{(j)}$ and the overcompleteness parameters $\nu^{(j)}$, which in turn determine the lattice parameter $(\bar{x}^{(j)}, \bar{k}_x^{(j)})$ via (1) and (5). As has been noted after (6), $\nu^{(j)}=0.5$ provides a convenient choice for the class of problems considered here. $b^{(j)}$ should be chosen according to the following criteria: (a) the ratio of the beamwidth $W_0^{(j)} = \sqrt{b^{(j)}/k}$ to the rough surface correlation length (see the discussion in Gordon *et al.*, 2005, Sec. V C). (b) the beamwidth should be large enough on the wavelength scale so that the propagating beams are well collimated and can be calculated locally along their propagation trajectories in the waveguide. Since $b^{(j)}$ represent the collimation distances of the beams, they should be chosen according to the spatial characteristics of the waveguide. To demonstrate these considerations in the homogeneous waveguide discussed here, we recall from Gordon *et al.*, 2005, Eq. (40) that the beamwidth is given roughly by $W(z_b) = \sqrt{b/k} \sqrt{1 + (z_b/b)^2}$, where z_b is the coordinate along the propagation trajectory, measured from the initiation point \mathbf{x}_m on the z_j boundary. Differentiating $\partial W(z_b)/\partial b$ for a given z_b , one finds that $W(z_b)$ is minimal if $b = z_b$. Thus, choosing $b^{(j)}$ roughly equal to or larger than the waveguide width $H = z_2 - z_1$ ensures that the footprint on the $z_{j'}$ plane of the beam that emanates from z_j is minimal.

G. Employing symmetries

In homogeneous media, the $\mathbf{S}^{(jj')}$ matrices have translation and left–right symmetries. Using the $\boldsymbol{\mu}=(\mathbf{m}, \mathbf{n})$ notations, they may be expressed as

$$S_{\mathbf{m}_j, \mathbf{n}_j, \mathbf{m}_{j'}, \mathbf{n}_{j'}}^{(jj')} = S_{(\mathbf{m}_j - \mathbf{m}_{j'}), \mathbf{n}_j, \mathbf{n}_{j'}}^{(jj')} = S_{-(\mathbf{m}_j - \mathbf{m}_{j'}), -\mathbf{n}_j, -\mathbf{n}_{j'}}^{(jj')} \quad (20)$$

This reduces the computational complexity of calculating \mathbf{S} , but does not reduce the *summation complexity*.

It should be noted that these symmetries still apply in vertically inhomogeneous media [though the matrix elements are different, as discussed, for example in (34)]. In range-dependent inhomogeneous configurations, on the other hand, the matrices lose these symmetries and the matrices need to be calculated for each initiation point and direction.

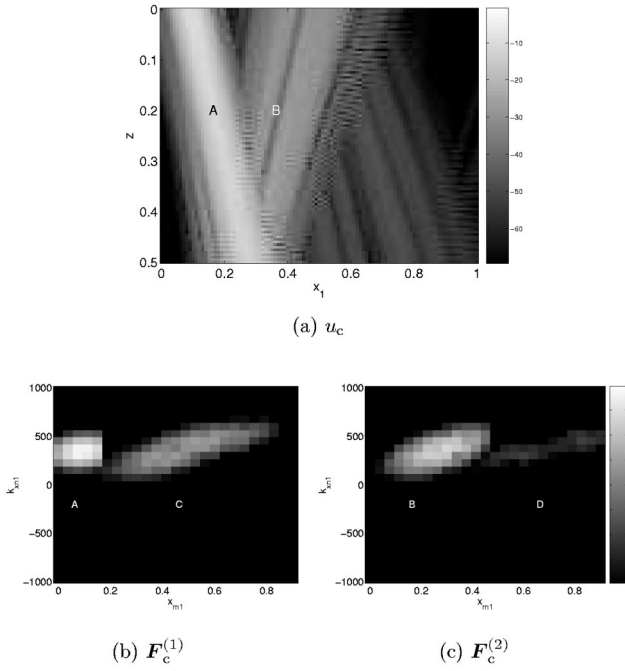


FIG. 2. The coherent field inside the waveguide. The parameters of the rough boundaries, of the illumination, and of the beam expansion are discussed in Sec. IV A. (a) The coherent intensity (dB) plotted in the (x_1, z) plane for $x_2=0$. Note the different scales of the x_1 and z axes. (b) and (c) The phase-space coefficient vectors (the beam amplitudes) $\mathbf{F}_c^{(1,2)}$ plotted in the (x_1, k_{x1}) phase space for $(x_2, k_{x2})=(0,0)$. Each point in the discrete lattice represents a single beam. The coefficient groups A, B, C, and D are associated with the corresponding fields in panel (a) as discussed in Sec. IV B.

IV. NUMERICAL EXAMPLE

The following example demonstrates the GBS method for the evaluation of the coherent and noncoherent fields and their phase-space characteristics within a doubly rough-surfaced waveguide.

A. Problem setup

Referring to Fig. 1, we consider a source-excited field in a homogeneous waveguide with wave number $k=1000$ and rough boundaries at $z_1=0$ and $z_2=H=0.5$ ($H \approx 80\lambda$), with Dirichlet boundary conditions. The source is a GB with collimation distance $b_T=1$ which emanates from $\mathbf{r}_T=(0,0,0.25)$ toward the surface z_1 at the initial direction $(\theta_T, \varphi_T)=(160^\circ, 0^\circ)$. Its field is thus given by

$$u^i(\mathbf{r}) = \frac{-ib_T}{z_b^i - ib_T} \exp\left[ik \left(z_b^i + \frac{1}{2} \frac{|\mathbf{x}_b^i|^2}{z_b^i - ib_T} \right) \right], \quad (21)$$

where (\mathbf{x}_b^i, z_b^i) are the incident beam's coordinates measured along the beam axis from the point \mathbf{r}_T . The beam waist at \mathbf{r}_T is $W_{0_T} = \sqrt{b_T/k} \approx 1/30$, and the far-zone diffraction angle is $\Theta_{D_T} = 1/\sqrt{kb_T} \approx (1/30)^\text{rad}$ [see Gordon *et al.*, 2005, Eqs. (40) and (73)].

For simplicity, the two rough surfaces have the same Gaussian spectra, with identical rms heights, $\sigma_{h1,2}$ and correlation lengths $l_{h1,2}$. They are given by $k\sigma_{h1,2} = 0.1$, implying that the scattering matrix can be calculated analytically

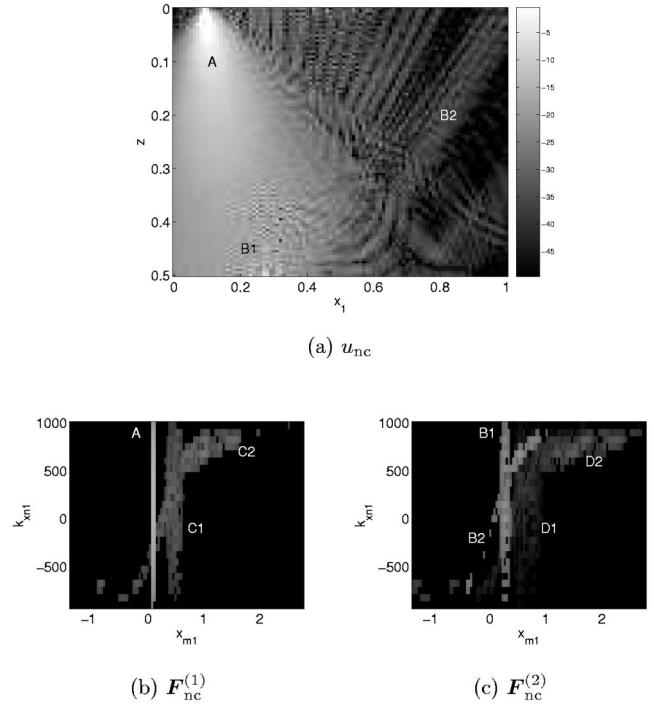


FIG. 3. Same as Fig. 2(a) for the noncoherent field and for the noncoherent phase-space coefficients $\mathbf{F}_{nc}^{(1,2)}$. The coefficient groups A, B, C, and D in panels (b), (c) are associated with the fields in panel (a) as discussed in Sec. IV C.

via the perturbation approach, and $l_{h1,2}^2 = 0.1W_{0_T}^2$, i.e., the correlation length is smaller than the incident beamwidth.

The collimation length of the expansion beams is now set to $b=0.5$, thus providing convenient GB propagators across the waveguide. This also facilitates the asymptotic evaluation of the projection integrals (9) for the GB2GB coherent and noncoherent scattering matrices.

The fields were calculated via the interactions series approach in Sec. III B with up to four successive reflections (the first one at z_1 , and then at z_2 , at z_1 , and at z_2 again). These reflections suffice in order to describe the field correctly up to the range $r=0.75 \approx 120\lambda$, beyond which the fields are weak. In Sec. V C we shall calculate the fields beyond this range up to $r=1.0$ in order to evaluate the bi-static reverberation as a function of the distance x_R between the source and the receiver, as described in Fig. 1.

B. Coherent field

The coherent intensity inside the waveguide, calculated via the GBS approach, is depicted in Fig. 2(a) in the plane of incidence (x_1, z) for $x_2=0$. This field has essentially the form of specular reflections. To gain further insight we also show in Figs. 2(b)–(c) the phase-space distribution of the beam coefficients $\mathbf{F}_c^{(1,2)}$ of (15) corresponding of the coherent fields. The plots in Figs. 2(b)–(c) depict slices of the four-dimensional phase space described by (x_1, k_{x1}) for $(x_2, k_{x2})=(0,0)$.

On each surface two non-negligible coefficient groups are readily discerned, denoted as A and C for $\mathbf{F}_c^{(1)}$ in Fig. 3(b) and B and D for $\mathbf{F}_c^{(2)}$ in Fig. 3(c). Group A corresponds to the first coherent reflections of the incident field from the

surface z_1 [i.e., the term $\mathbf{S}_c^{(0)}$ in (14)], group B corresponds to the second coherent reflections from z_2 [i.e., the term $\mathbf{S}_c^{(0)}\mathbf{S}_c^{(12)}$ in (14) corresponding to scattering at z_1 and then at z_2], while groups C and D correspond to the third and fourth coherent reflections from z_1 and then from z_2 , respectively (the terms $\mathbf{S}_c^{(0)}\mathbf{S}_c^{(12)}\mathbf{S}_c^{(21)}$ and $\mathbf{S}_c^{(0)}\mathbf{S}_c^{(12)}\mathbf{S}_c^{(21)}\mathbf{S}_c^{(12)}$ there). These phase-space distributions demonstrate the loss, spread, and advance of the specular reflections as the range is increased. Each successive interaction with the surfaces is weaker due to the propagation loss between the surfaces and the loss of energy to the noncoherent component [i.e., $\bar{V} < 1$ in (9a)]. Note also that the phase-space distributions are spatially wider, mainly due to the propagation spread of the coherent beam, but the spectral spread is essentially unchanged. These characteristics can also be seen from the field results in Fig. 2(a), where one may also discern the interference patterns between the coherently reflected fields.

C. Noncoherent field

The noncoherent intensity inside the waveguide is shown in Fig. 3(a). Unlike the coherent field, here the field is wide and has a rather intricate structure. To understand these features we refer to Figs. 3(b)–(c) for the phase-space coefficients $\mathbf{F}_{nc}^{(1,2)}$ at the $z_{1,2}$ surfaces, respectively. The presentation formats in these figures is the same as in Fig. 2.

Several coefficient groups are readily discerned in these figures. Group A in $\mathbf{F}_{nc}^{(1)}$ corresponds to the first noncoherent interaction of the incident field with z_1 , as described by $\mathbf{S}_{nc}^{(0)}$ in (17a). As explained in Gordon *et al.*, 2005, Fig. 5(c), it is localized about the center of mass of the incident beam and consists of noncoherent reflections from that point in all directions; hence, it is described by spectrally wide vertical trace in the phase space. These contributions are readily observed in terms of the fan beam in Fig. 3(a) that emerges from the center of mass of the incident beam (point $x_1 = 1$ on the z_1 surface).

Groups B1, B2 in $\mathbf{F}_{nc}^{(2)}$ correspond to the second internal scattering (i.e., scattering at z_1 and then at z_2). Group B1 is the noncoherent scattering from z_2 of the coherent scattered beams that propagated from z_1 [i.e., $\mathbf{S}_c^{(0)}\mathbf{S}_{nc}^{(12)}$ in (17b), also denoted briefly as $c\text{-}nc$]. This group has the typical noncoherent footprint of a vertical trace in the phase space, though it is centered now around the center of mass of the coherent reflection from z_1 (i.e., around $x_1 = 0.27$ there), and is wider due to the spreading of the incident beam. In Fig. 3(a), the resulting field is a fan beam emerging from $x_1 = 0.27$. Group B2 is the coherent scattering at z_2 of the noncoherent beams that propagated from z_1 [i.e., $\mathbf{S}_{nc}^{(0)}\mathbf{S}_c^{(12)}$ in (17b), also denoted $nc\text{-}c$]. The diagonal phase-space footprint is due to the fact that the noncoherent scattered beams from z_1 hit surface z_2 at points that depend on their exit directions from z_1 , and are then coherently reflected. This contribution is seen in Fig. 3(a) as the spread beams that originate from different points on the z_2 plane and at correspondingly different directions (this contribution can be visualized as a specular reflection of fan beam A at z_2). As discussed in Sec. III D, the noncoherent scattering of the noncoherent beams from the upper surface is neglected.

The coefficient group of the third internal scattering consists of two groups (C1, C2 in $\mathbf{F}_{nc}^{(1)}$), whose footprints are similar to group B. C1 represents the $c\text{-}c\text{-}nc$ scattering and its footprint is similar to B1 in $\mathbf{F}_{nc}^{(2)}$, except that here the vertical trace is farther away (around $x_1 = 0.45$) and wider (due to the beam spread of the incident coherent field). C2 represents the $nc\text{-}c\text{-}c$ and $c\text{-}nc\text{-}c$ scattering (the footprints of these two scattering events are almost indistinguishable and they are therefore considered as a single group). It is of the same type as B2, where the diagonal phase-space trace is due to the relationship between the exit point and the direction of the coherent specular reflection at z_1 . These contributions are already too weak to be noticed in Fig. 3(a).

Finally, the coefficient group of the fourth internal scattering consists also of two groups: D1 with vertical trace represents the $c\text{-}c\text{-}c\text{-}nc$ interaction and is similar to B1 and C1, while D2 is similar to B2 and C2 and is composed of $nc\text{-}c\text{-}c\text{-}c$, $c\text{-}nc\text{-}c\text{-}c$, and $c\text{-}c\text{-}nc\text{-}c$ interactions. Again, these contributions are too weak to be noticed in Fig. 3(a).

V. GBS REPRESENTATION OF THE RECEIVER PATTERN

The theory in Sec. III has been phrased for field calculations. In this section we would like to extend the theory so that it now includes the coupling between the multiply reflected beam propagators and the receiving antenna which is characterized by a receiving pattern.

One useful implementation of the aforementioned technique is the evaluation of bistatic reverberations in underwater sounding or communication systems. The problem setup is depicted in Fig. 1. A transmitter T with a given radiation pattern is positioned at (\mathbf{x}_T, z_T) and transmits (upward) within a doubly rough surface homogeneous waveguide with upper and lower boundaries at $z_{1,2}$, respectively. The receiver R with a given pattern $f_R(\theta, \phi)$ is located at $\mathbf{r}_R = (\mathbf{x}_R, z_R)$.

A. The receiver model

From antenna theory (Collin and Zucker, 1969), the signal at a receiving antenna when located in the origin is given by

$$R = C \int_{4\pi} \sin \theta d\theta d\phi f_R(\theta, \phi) \tilde{u}^i(\theta, \phi), \quad (22)$$

where the integral encompasses all directions of incidence with $\tilde{u}^i(\theta, \phi)$ being the spectral amplitude of the incidence field with respect to the (θ, ϕ) direction [normalized such that for an incident plane wave $u^i(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}}$, with $\mathbf{k}^i = -k(\sin \theta^i \cos \phi^i, \sin \theta^i \sin \phi^i, \cos \theta^i)$ one has $\tilde{u}^i(\theta, \phi) = (1/\sin \theta) \delta(\theta - \theta^i) \delta(\phi - \phi^i)$]. In (22), $f_R(\theta, \phi)$ is the radiation pattern of antenna R into the (θ, ϕ) direction once this antenna operates in “radiation mode,” normalized such that the field radiated by antenna R due to a unit excitation is

$$u_R^t(\mathbf{r}) = \frac{e^{i\mathbf{k} \cdot \mathbf{r}}}{4\pi|\mathbf{r}|} f_R(\theta, \phi). \quad (23)$$

Under these conditions, it follows via the reciprocity theorem that C in (22) is a universal proportionality constant (Collin and Zucker, 1969).

In order to embed the receiver within the GBS formulation, we need to express the spectral representation (22) within a Cartesian spectrum formulation. Thus, we note from (23) that the field radiated by this antenna, when located in the origin, can be expressed as

$$u_{\mathbf{R}}^i(\mathbf{r}) = \frac{1}{(2\pi)^2} \int d^2k_x \frac{1}{-2ik_z} \tilde{f}_{\mathbf{R}}^{\pm}(\mathbf{k}_x) e^{i\mathbf{k}_x \cdot \mathbf{x} + ik_z|z|}, \quad (24)$$

where upper and lower signs correspond to an observation point at $z \geq 0$, respectively, $k_z = \sqrt{k^2 - |\mathbf{k}_x|^2}$ with $\text{Im } k_z \geq 0$, and

$$\tilde{f}_{\mathbf{R}}^{\pm}(\mathbf{k}_x) = f_{\mathbf{R}}(\theta, \phi)|_{k_x = k \sin \theta (\cos \phi, \sin \phi)}, \quad (25)$$

where the \pm is for $0 < \theta < \pi/2$ and $\pi/2 < \theta < \pi$, respectively. Equation (22) can therefore be expressed instead as

$$R = \frac{C}{(2\pi)^2} \int d^2k_x [\tilde{f}_{\mathbf{R}}^{-}(-\mathbf{k}_x) \tilde{u}^{i+}(\mathbf{k}_x) + \tilde{f}_{\mathbf{R}}^{+}(-\mathbf{k}_x) \tilde{u}^{i-}(\mathbf{k}_x)], \quad (26)$$

where $\tilde{u}^{i\pm}(\mathbf{k})$ are the $\pm z$ propagating spectral components of the incident field u^i at the antenna plane $z=0$.

Finally, let the receiver be located in the waveguide at $\mathbf{r}_{\mathbf{R}} = (\mathbf{x}_{\mathbf{R}}, z_{\mathbf{R}})$, $z_1 \leq z_{\mathbf{R}} \leq z_2$. We also express the $\pm z$ -propagating spectra in terms of their values $\tilde{u}_1^{+}(\mathbf{k}_x)$ and $\tilde{u}_2^{-}(\mathbf{k}_x)$ at $z_{1,2}$, respectively. Then (26) becomes

$$R(\mathbf{r}_{\mathbf{R}}) = \frac{C}{(2\pi)^2} \int d^2k_x [\tilde{f}^{-}(-\mathbf{k}_x) \tilde{u}_1^{+}(\mathbf{k}_x) e^{ik_z(z_{\mathbf{R}} - z_1)} + \tilde{f}^{+}(-\mathbf{k}_x) \tilde{u}_2^{-}(\mathbf{k}_x) e^{ik_z(z_2 - z_{\mathbf{R}})}] e^{i\mathbf{k}_x \cdot \mathbf{x}_{\mathbf{R}}}. \quad (27)$$

B. GBS formulation

Next, we apply the frame decomposition and the Gaussian beam summation formulation to describe the downward/upward propagating waves in (27). Utilizing the matrix notation for brevity, one finds from (15), ignoring the u^i term, that

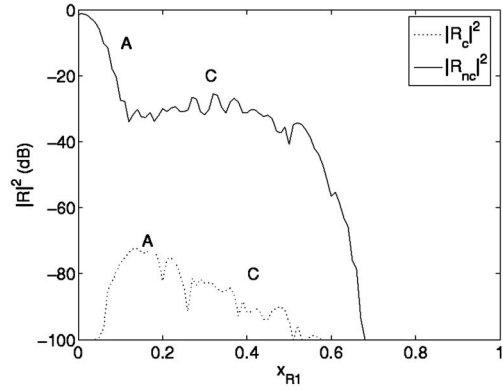
$$u_1^{+}(\mathbf{r})|_{z_1} = \mathbf{F}^{(1)} \boldsymbol{\Psi}^{(1)}(\mathbf{x}), \quad u_2^{-}(\mathbf{r})|_{z_2} = \mathbf{F}^{(2)} \boldsymbol{\Psi}^{(2)}(\mathbf{x}), \quad (28)$$

where we use the frame element vectors $\boldsymbol{\Psi}^{(j)}(\mathbf{x}) = \psi_{\boldsymbol{\mu}_j}^{(j)}(\mathbf{x})$ and the coefficients row vectors $\mathbf{F}^{(j)}$. Substituting into (27), one finds that $R(\mathbf{r}_{\mathbf{R}})$ has the form in (15) with the u^i term ignored, except that the propagator vectors $\mathbf{B}^{(j)}$ are now replaced by the received (or measured) beam vectors $\mathbf{R}^{(j)}$, i.e.,

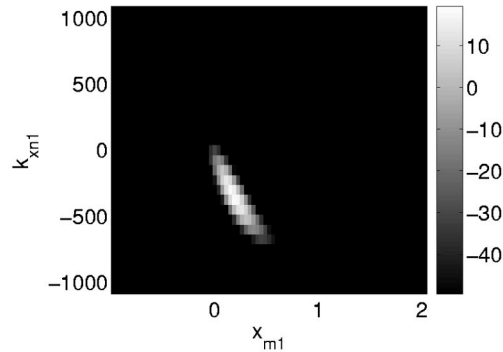
$$R_{\boldsymbol{\mu}}(\mathbf{r}_{\mathbf{R}}) = \mathbf{F}^{(1)} \mathbf{R}^{(1)}(\mathbf{r}_{\mathbf{R}}) + \mathbf{F}^{(2)} \mathbf{R}^{(2)}(\mathbf{r}_{\mathbf{R}}), \quad (29)$$

where

$$R_{\boldsymbol{\mu}_{1,2}}^{(1,2)}(\mathbf{r}_{\mathbf{R}}) = \frac{C}{(2\pi)^2} \int d^2k_x \tilde{\psi}_{\boldsymbol{\mu}_{1,2}}^{(1,2)}(\mathbf{k}_x) \times e^{i\mathbf{k}_x \cdot \mathbf{x}_{\mathbf{R}} + ik_z|z_{\mathbf{R}} - z_{1,2}|} \tilde{f}_{\mathbf{R}}^{\pm}(-\mathbf{k}_x). \quad (30)$$



(a) $|R_{c,nc}|^2$



(b) $\mathbf{R}^{(1)}(0)$

FIG. 4. Bistatic reverberations in double rough surface waveguide, for the system configuration depicted in Fig. 1 and the problem parameters in Sec. V C. (a) The coherent and noncoherent reception intensity $|R_c|^2$ and $|R_{nc}|^2$ (in dB) as a function of the range x_{R1} from the transmitter to the receiver. (b) The receiver's phase-space vector $\mathbf{R}^{(1)}(0)$ of (30).

Note from Gordon *et al.*, 2005, Eq. (29), and that for $C = \tilde{f}_{\mathbf{R}}^{\pm} = 1$ these integrals define the beam propagator contributions at $\mathbf{r}_{\mathbf{R}}$, i.e., $R_{\boldsymbol{\mu}_j}^{(j)}(\mathbf{r}_{\mathbf{R}}) \rightarrow B_{\boldsymbol{\mu}_j}^{(j)}(\mathbf{r}_{\mathbf{R}})$. Thus, if $\tilde{f}^{r\pm}(\mathbf{k}_x)$ are slowly varying functions of \mathbf{k}_x relative to $\tilde{\psi}(\mathbf{k}_x)$, one finds via asymptotic evaluation that

$$R_{\boldsymbol{\mu}_{1,2}}^{(1,2)}(\mathbf{r}_{\mathbf{R}}) = C \tilde{f}_{\mathbf{R}}^{\pm}(-\mathbf{k}_{x\mathbf{n}}) \mathbf{B}_{\boldsymbol{\mu}_{1,2}}^{(1,2)}(\mathbf{r}_{\mathbf{R}}). \quad (31)$$

C. Numerical example: Bistatic reverberations

As a demonstration of the receiver's pattern effect, we shall calculate the bistatic reverberation in the two antenna system depicted in Fig. 1 as a function of the range from the transmitter T to the receiver R which is located in the plane of incidence $x_2=0$ at $\mathbf{r}_{\mathbf{R}} = (x_{R1}, 0, 0.25)$, where x_{R1} is the range. The pattern of R is Gaussian with a maximum in the $(\theta_{\mathbf{R}}, \phi_{\mathbf{R}}) = (160^\circ, 0^\circ)$ direction (i.e., it is aimed in the plane of incidence toward surface z_1 and away from T), and with beamwidth $\Theta_{\mathbf{R}} = 1/\sqrt{k b_{\mathbf{R}}} = (1/10)^{\text{rad}}$, where we used $b_{\mathbf{R}} = 0.1$. The beam is oblique but stigmatic (circular symmetric), and therefore we find from (25) that $\tilde{f}_{\mathbf{R}}^{+} = 0$ and

$$\tilde{f}_{\mathbf{R}}^{-}(\mathbf{k}_x) = \sqrt{2b_{\mathbf{R}}k/\cos \theta_{\mathbf{R}}} e^{-i(\mathbf{k}_x - \mathbf{k}_{\mathbf{R}})^T \boldsymbol{\Gamma}_{\mathbf{R}}(\mathbf{k}_x - \mathbf{k}_{\mathbf{R}})/2k}, \quad (32)$$

where $\mathbf{k}_{\mathbf{R}} = k \sin \theta_{\mathbf{R}} (\cos \phi_{\mathbf{R}}, \sin \phi_{\mathbf{R}})$, $\boldsymbol{\Gamma}_{\mathbf{R}}$ is a diagonal matrix with $\Gamma_{R,11} = b_{\mathbf{R}}/\cos^2 \theta_{\mathbf{R}}$ and $\Gamma_{R,22} = b_{\mathbf{R}}$, and the superscript T

in the exponent denotes a transposed vector. We have also used the gain-oriented normalization:

$$\frac{1}{2\pi} \int |\tilde{f}_R^-(\theta, \phi)|^2 d\Omega = 1. \quad (33)$$

Since $b_R \ll b$ here (i.e., the antenna pattern is wider than the spectral beamwidth), we may approximate $R_{\mu_1}^{(1)}(\mathbf{r}_R)$ by (31) with $\tilde{f}_R^-(\mathbf{k}_x)$ from (32). The other physical parameters are the same as in the example of Sec. IV.

The received coherent and noncoherent intensities $|R_c|^2$ and $|R_{nc}|^2$ as a function of the range x_{R1} are shown in Fig. 4(a). As can be seen, $|R_{nc}|$ is stronger and wider than $|R_c|$. This can be explained by noting that the receiving pattern of R is sensitive to waves that are scattered from surface z_1 and propagate backward towards T. Such components exist mainly in the noncoherent field.

To quantify this observation we refer to Fig. 4(b), which depicts the phase-space distribution of the receiver's vector $\mathbf{R}^{(1)}$ of (30) ($\mathbf{R}^{(2)}$ is negligible) at $\mathbf{r}_R = (0, 0, 0.25)$. As expected, $\mathbf{R}^{(1)}$ is centered spatially about $x_1 = x_{R1} - 0.25 \tan \theta_R$, which is the footprint of R at the z_1 plane, and spectrally about $k_{x1} = -k \sin \theta_R$. As x_{R1} changes, the footprints in Fig. 4(b) also move horizontally.

Now, following (29), the received signals $|R_c|$ and $|R_{nc}|$ are obtained by superimposing the phase-space distribution in Fig. 4(b) onto the distributions in Figs. 2(b)–(c) and 3(b)–(c). In other words, $\mathbf{R}^{(1)}$ enhances only those beams in Figs. 2(b)–(c) and 3(b)–(c) that fall within its footprint in Fig. 4(b). Thus, for $|R_c|$ one observes that none of the coefficient groups in Fig. 2(b) falls under the footprint of $\mathbf{R}^{(1)}$ in Fig. 4(b). One obtains, however, some weak contributions in the near zone from coefficient groups A and C. For $|R_{nc}|$, one readily observes two main contributions from coefficient groups A and C in Fig. 3(b).

VI. FUTURE WORK: EXTENSION TO INHOMOGENEOUS MEDIA

The GBs can be propagated in inhomogeneous media or through curved interfaces, where they retain their Gaussian functional form throughout the propagation process (see Babich and Popov, 1989; Popov, 1982; Porter and Bucker, 1987; Keenan and Weinberg, 2001; Heyman and Felsen, 2001; Nowack, 2003). Thus, the frame-based GBS representation for propagation in rough surface waveguide can readily be implemented in such media. One should distinguish though between range-independent media (purely vertically stratified) and range-dependent configurations. In both cases we use the uniform phase-space lattice discussed in Sec. III A, yet in the former, the matrix symmetries in (20) still apply [though the matrix elements are different, as discussed, for example in (34)], whereas in the latter there are no such symmetries and the matrices need to be calculated for each initiation point and direction. Since the beams and the scattering matrices are still localized, the \mathbf{S} matrices are sparse and one may still use the beam-based localization procedures discussed in Secs. III D and III E. The inhomogeneous media can create a situation depicted in Fig. 5 where a scattered beam is refracted and interacts with the same rough

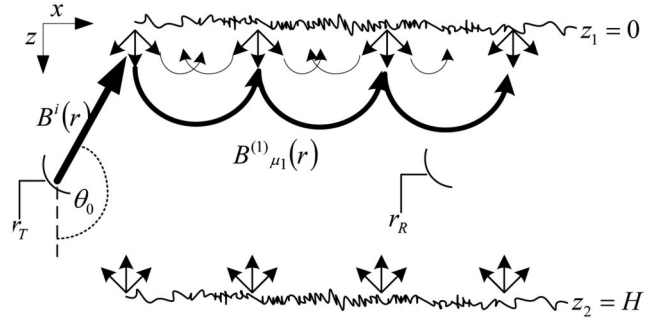


FIG. 5. The GBS setting in a doubly rough boundary nonhomogeneous waveguide which supports trapped fields near surface z_1 . The spectral constituents associated with the trapped field are described by the scattered beam emerging from the rough surface z_1 and refracting back to the same surface without interacting with the other boundary [see the scattering matrix $\mathbf{S}^{(11)}$ in (35)].

boundary it had originated from without reaching the other boundary (e.g., surface ducts). This situation is easily taken care of by adding two terms to Eqs. (16a)–(16b)

$$\mathbf{F}^{(1)} = \mathbf{S}^{(0)} + \mathbf{S}^{(21)}\mathbf{F}^{(2)} + \mathbf{S}^{(11)}\mathbf{F}^{(1)}, \quad (34a)$$

$$\mathbf{F}^{(2)} = \mathbf{S}^{(12)}\mathbf{F}^{(1)} + \mathbf{S}^{(22)}\mathbf{F}^{(2)}, \quad (34b)$$

where

$$\mathbf{S}^{(11)} = S(B_{\mu_1}^{(1)}|_{z_1}, \psi_{\mu_1}^{(1)}), \quad (35a)$$

$$\mathbf{S}^{(22)} = S(B_{\mu_2}^{(2)}|_{z_2}, \psi_{\mu_2}^{(2)}). \quad (35b)$$

Note that the $\mathbf{S}^{(11)}$ has nonvanishing elements only for those μ_1 for which the elements of $\mathbf{S}^{(12)}$ vanish, i.e., each beam originating from the upper surface either reaches the lower surface or refracts to the upper surface.

The equation set (34) may be set in the form

$$\begin{pmatrix} \mathbf{I} - \mathbf{S}^{(11)} & -\mathbf{S}^{(21)} \\ -\mathbf{S}^{(12)} & \mathbf{I} - \mathbf{S}^{(22)} \end{pmatrix} \begin{pmatrix} \mathbf{F}^{(1)} \\ \mathbf{F}^{(2)} \end{pmatrix} = \begin{pmatrix} \mathbf{S}^{(0)} \\ \mathbf{0} \end{pmatrix}, \quad (36)$$

and then solved either via successive approximations as in (17), or collectively via matrix inversion as in (18).

VII. CONCLUDING REMARKS

We introduced an extension of the GBS formulation of rough surface scattering to doubly rough surface waveguide scenario. The theory utilizes deterministic GB propagators and stochastic GB2GB scattering matrices to describe multiple scattering from the upper and lower rough boundaries. The resulting coherent and noncoherent fields were shown to be fully described by the phase-space beam amplitude vectors $\mathbf{F}^{(1,2)}$. Two alternative representations for $\mathbf{F}^{(1,2)}$ were presented: a series of successive interactions and a collective solution (Secs. III B and III C, respectively). As has been demonstrated by the numerical examples in Secs. IV and V C, these amplitude vectors reveal explicitly the local phase-space footprint of the stochastic scattering process, thus providing a cogent physical interpretation to the field propagation. These localization features were utilized to obtain an effective mathematical algorithm, as discussed in Secs. III D–III F. As discussed there, controlling the expan-

sion parameters (e.g., the beam collimation) changes the expansion characteristics continuously from a ray-like solution to a plane-wave-like solution; hence, these parameters should be selected properly to match a given propagation configuration.

In Sec. V the phase-space beam formulation has been extended to accommodate the reception by an antenna with a receiving pattern. An implementation to bistatic reverberations inside a rough surface waveguide as a function of the range and of the source and the receiver directions has been analyzed as an example. The GBS results should be compared with the spectral perturbation approach that was used in Schmidt and Kuperman (1995) to study reverberations in stratified waveguides with rough surface boundaries within the shallow angles and low frequencies regime.

Finally, a simple extension to inhomogeneous media has been discussed qualitatively in Sec. VI, where we mainly considered the formulation of possible surface ducts near the rough surface boundaries. Further work is needed to explore the full power of this elegant technique.

ACKNOWLEDGMENTS

This paper is dedicated to the memory of our coauthor, the late Professor Reuven Mazar. E. Heyman acknowledges partial support by the Israel Science Foundation under Grant No. 216/02 and by the US–Israel Binational Science Foundation, Jerusalem, Israel, under Grant No. 9900448.

Babich, V. M., and Popov, M. M. (1989). “Gaussian summation method (review),” *Radiophys. Quantum Electron.* **39**, 1063–1081.
 Beenakker, C. W. J. (1997). “Random-matrix theory of quantum transport,” *Rev. Mod. Phys.* **69**(3), 731–808.

Collin, R. E., and Zucker, F. J. (1969). *Antenna Theory* (McGraw-Hill, New York), Vol. 1, Chap. 4.
 Etter, P. C. (2001). “Recent advances in underwater acoustic modeling and simulation,” *J. Sound Vib.* **240**(2), 351–383.
 Etter, P. C. (2003). *Underwater Acoustics Modeling and Simulations*, 3rd ed. (Spon, New York).
 Gordon, G., Heyman, E., and Mazar, R. (2005). “A phase-space Gaussian beam summation representation of rough surface scattering,” *J. Acoust. Soc. Am.* **117**, 1911–1921.
 Heyman, E., and Felsen, L. B. (2001). “Gaussian beam and pulsed beam dynamics: Complex source and spectrum formulations within and beyond paraxial asymptotics,” *J. Opt. Soc. Am. A* **18**, 1588–1611.
 Keenan, R. E., and Weinberg, H. (2001). “Gaussian ray bundle (GRAB) model shallow water acoustic workshop implementation,” *J. Comput. Acoust.* **9**, 133–148.
 Mazar, R., and Felsen, L. B. (1989). “Stochastic geometrical theory of diffraction (SGTD),” *J. Acoust. Soc. Am.* **86**(6), 2292–2308.
 Nowack, R. L. (2003). “Calculation of synthetic seismograms with Gaussian beams,” *Pure Appl. Geophys.* **160**, 487–507.
 Popov, M. M. (1982). “A new method of computation of wave fields using Gaussian beams,” *Wave Motion* **4**, 85–97.
 Porter, M. B., and Bucker, H. P. (1987). “Gaussian beam tracing for computing ocean acoustic fields,” *J. Acoust. Soc. Am.* **82**, 1349–1359.
 Schmidt, H., and Kuperman, W. A. (1995). “Spectral representations of rough interface reverberation in stratified ocean waveguides,” *J. Acoust. Soc. Am.* **97**, 2199–2209.
 Sheng, P. (ed). (1990). *Scattering and Localization of Classical Waves in Random Media* (World Scientific, Singapore).
 Shlivinski, A., Heyman, E., Boag, A., and Letrou, C. (2004). “A phase-space beam summation formulation for wideband radiation,” *IEEE Trans. Antennas Propag.* **52**, 2042–2056.
 Snchez-Gil, J. A., Freilikher, V., Yurkevich, I., and Maradudin, A. A. (1998). “Coexistence of ballistic transport, diffusion, and localization in surface disordered waveguides,” *Phys. Rev. Lett.* **80**(5), 948–951.
 Voronovich, A. (1999). *Wave Scattering from Rough Surfaces* (Springer, Berlin).

Geoacoustic inversion in time domain using ship of opportunity noise recorded on a horizontal towed array

Cheolsoo Park

*Korea Ocean Research & Development Institute/Korea Research Institute of Ships & Ocean Engineering,
171 Jang-Dong Yuseong, Daejeon, Korea*

Woojae Seong

Department of Naval Architecture and Ocean Engineering, Seoul National University, Seoul 151-744, Korea

Peter Gerstoft

*Marine Physical Laboratory, Scripps Institute of Oceanography, University of California,
La Jolla, California 92093*

(Received 6 September 2004; revised 6 January 2005; accepted 8 January 2005)

A time domain geoacoustic inversion method using ship noise received on a towed horizontal array is presented. The received signal, containing ship noise as a source of opportunity, is time-reversed and then back-propagated to the vicinity of the ship. The back-propagated signal is correlated with the modeled signal which is expected to peak at the ship's location in case of a good match for the environment. This match is utilized for the geoacoustic parameter inversion. The objective function for this optimization problem is thus defined as the normalized power focused in an area around the source position, using a matched impulse response filter. A hybrid use of global and local search algorithms, i.e., GA and Powell's method is applied to the optimization problem. Applications of the proposed inversion method to MAPEX 2000 noise experiment conducted north of the island of Elba show promising results, and it is shown that the time domain inversion takes advantage of dominant frequencies of the source signature automatically. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1862574]

PACS numbers: 43.30.Pc [AIT]

Pages: 1933–1941

I. INTRODUCTION

Nowadays the use of passive sonar systems to acquire the information of sources, especially in shallow water areas, is increasing. In addition, various numerical models have been used to simulate their performances with measured or estimated environments, among which the seabed properties are considered to be one of the most influential parameters on the shallow water sound propagation. Therefore, various techniques to estimate the geoacoustic parameters of the ocean bottom via remote sensing have been developed in order to resolve the problem of high cost and time consumption in relation to direct measurement of the sub-bottom properties.

Most areas of the ocean, especially in shallow waters, show varying seabed characteristics that require range-dependent parametrizations for geoacoustic inversion. However, a full range-dependent inversion that can identify the ocean bottom structure simultaneously has not yet been achieved due to computational burdens in forward modeling. A majority of the inversions in underwater acoustics have been performed with the assumption of range-independent environments using long-range propagation data received on a vertical line array (VLA) that produces spatially averaged output.^{1,2} In order to avoid this degradation of resolution, towed horizontal line array (HLA) inversion schemes have been proposed and implemented.^{3–6} The short distance between a sound source and HLA enables us to simplify the range dependence into local segments of range-independent sectors.⁶ The towed HLA system possesses additional advan-

tages of mobility and manageability, which are inherently suitable for a towed sonar system compared to a VLA.

Inversions which utilize the full structure of the bottom reflected signals usually require well controlled acoustic sources. Such dedicated sources may add extra burdens to an inversion experiment especially of the towed HLA system. Contrary to controlled sources, the noise of the tow ship or ship of opportunity can be a good alternative source for inversion.^{5,7} Most of the analyses with the acquired data are processed in the frequency domain, both at the narrowband and broadband, with the exception of a few that have been chosen to deal directly in the time domain.^{6,8–13} Since the ship noise shows a broadband characteristic, it is needed to extract strong tonal components in frequency domain inversions.⁵ On the contrary, time domain inversions can take advantage of a full broadband signal without having to go through the necessary processing.

The inversion is usually formulated as an optimization problem enabling various optimization techniques to be applied to the inverse problems, such as a global optimization^{14,15} and a hybrid form of global and local search.^{16,17} The solutions of the optimization can be regarded as the parameters which optimize the objective function defined to measure the difference (or similarity) between received and simulated data. Although it is common for all parameters to be inverted simultaneously, iterative optimizations can be performed after grouping the parameters based on their sensitivities.¹⁸

In this paper, a time domain inversion using ship noise

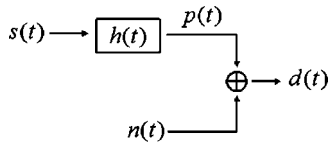


FIG. 1. An idealized data acquisition system where the received signal $d(t)$ is comprised of the propagated signal $p(t)$ and noise $n(t)$.

acquired on a towed HLA is proposed. An objective function is defined using a matched impulse response filter for the unknown source signal and the filtered output is implemented by back-propagating the time-reversed signal using a ray-based model. Then, a series of global and local searches is applied to measured data in order to find the optimum parameters of the inverse problem.

The paper is organized as follows. In Sec. II we describe a matched impulse response filter which is a basic theory used in defining the objective function. In Sec. III, a sensitivity test is also presented related to the objective function. In Sec. IV we describe the MAPEX 2000 noise experiment and the parametrization of the inversion. In Sec. V we present the geoacoustic inversion results using the source of opportunity.

II. MATCHED IMPULSE RESPONSE FILTER

It has been demonstrated that the sound wave received by a hydrophone and retransmitted backwards focuses at the source position.¹⁹ This phenomenon can be implemented numerically using a received signal as a virtual source, which was utilized in source localization.^{20,21} It has also been used in underwater applications such as pulse compressing for the communications in the name of time reversal or phase conjugation.^{22,23} The wave focusing is explained via matched filtering and array processing.

A data acquisition process can be represented by an idealized system of Fig. 1, where the received signal $d(t)$ is comprised of propagated signal $p(t)$ and noise $n(t)$. In the linear time invariant system, the propagation of source signal $s(t)$ through a medium can be simply expressed by convolution with an impulse response $h(t)$ as

$$d(t) = p(t) + n(t) = s(t) * h(t) + n(t). \quad (1)$$

The rigorous impulse response for wave propagating in the ocean can be found by solving the wave equation. However, for a set of source and receiver separated within a few water depths, the received signal can be approximated by the sum of eigen-rays weighted and delayed according to their propagation paths. Thus, ignoring the noise for the moment, the received signal can be expressed as

$$d(t) = s(t) * h(t) = s(t) * \sum_{i=1}^M a_i \delta(t - \tau_i), \quad (2)$$

where a_i and τ_i represent the amplitude decay and travel time of the signal component which are propagated through M different paths (eigen-rays), respectively.

If this signal is time reversed and back-propagated, then the medium impulse response simply becomes $h(-t)$, called

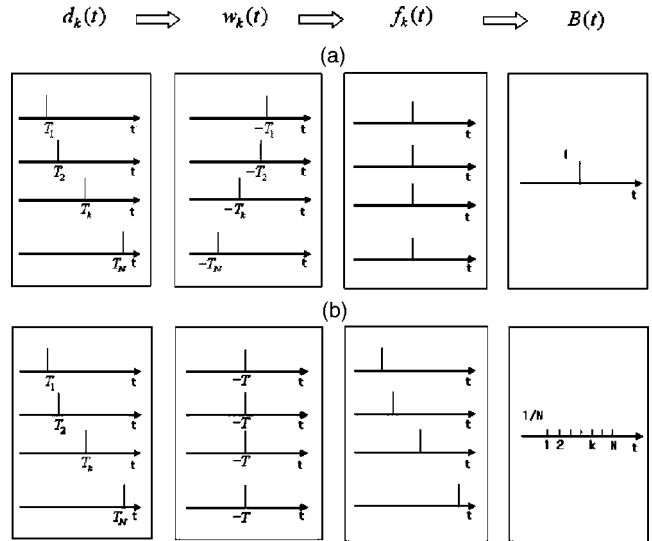


FIG. 2. Examples of beamforming using an impulse signal; (a) optimally beamformed signal and (b) nonoptimally beamformed signal.

the matched impulse response. Applying this as a filter $w(t)$ produces an output at the source location given by

$$f(t) = d(t) * h(-t) = \sum_{i=1}^M a_i^2 s(t) + \sum_{i=1}^M \sum_{j \neq i}^M a_i a_j s(t - (\tau_i - \tau_j)). \quad (3)$$

It can be seen from Eq. (3) that the filter output comprises of two components, $\sum_{i=1}^M a_i^2 s(t)$ and $\sum_{i=1}^M \sum_{j \neq i}^M a_i a_j s(t - (\tau_i - \tau_j))$, which correspond to a mainlobe and multiple sidelobes of the signal, respectively.

Now, the plane-wave beam-forming concept can be extended for an array with N receivers such that the back-propagated output power is defined as

$$B(t) = \frac{1}{N} \sum_{k=1}^N f_k(t) = \frac{1}{N} \sum_{k=1}^N \sum_{i=1}^M a_{i,k}^2 s(t) + \frac{1}{N} \sum_{k=1}^N \sum_{i=1}^M \sum_{j \neq i}^M a_{i,k} a_{j,k} s(t - (\tau_{i,k} - \tau_{j,k})). \quad (4)$$

If the filter is correct, the mainlobe will be reinforced whereas the sidelobes are cancelled by destructive interference. As a result, the mainlobe becomes dominant over the sidelobes as the number of channels increase [Fig. 2(a)]. If the filter is presumed incorrectly, which corresponds to an environmental mismatch, there will be multiple mainlobes along with additional sidelobes [Fig. 2(b)].

The ocean environment itself can be considered as a filter when the received signal is retransmitted reversely. Therefore, when the environment and also the source receiver configurations are modeled as when the signal was recorded at the receiver array, the retransmitted signal will be focused at the original source position.

III. GEOACOUSTIC INVERSION USING SHIP NOISE

Acoustic signals transmitted by a source convey information of the medium through which they propagated. There are numerous acoustic sources for ocean experiments, and the noise generated by a ship is one which will be present most of the time when experiments are carried out in the ocean. If the waveform of a ship noise is unknown, only the impulse response can be used in the geoacoustic inversion.

Inversion parameters which describe the environment are represented by a model vector $\bar{\mathbf{m}} = [\bar{m}_1, \bar{m}_2, \dots, \bar{m}_L]^T$, where $[]^T$ is a transpose operator. The normalized filter output of the channel is defined as

$$f_k(t, \bar{\mathbf{m}}) = \frac{d_k(T-t)}{\sqrt{\int_{-\infty}^{\infty} |d_k(T-t)|^2 dt}} * \frac{\bar{h}_k(t, \bar{\mathbf{m}})}{\sqrt{\int_{-\infty}^{\infty} |\bar{h}_k(t, \bar{\mathbf{m}})|^2 dt}}, \quad (5)$$

where $\bar{h}_k(t, \bar{\mathbf{m}})$ is an impulse response of the corresponding environment. Equation (5) is implemented by propagating the received and time-reversed signal through the environment using a forward model. The objective function to be

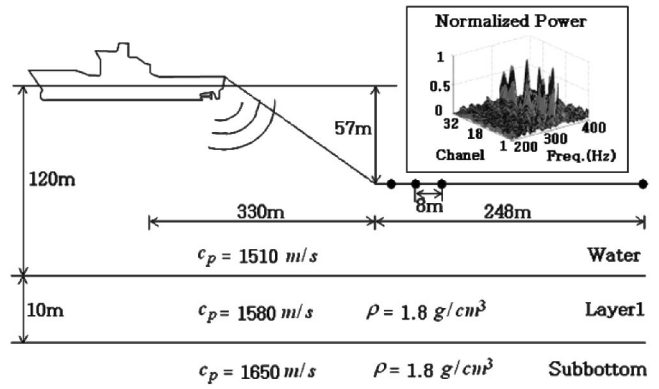


FIG. 3. The shallow water environment and the geometry of towed array system used for the sensitivity test. The power spectra of measured signals which are simulated based on the given environment are also shown in the figure.

maximized for the inversion using N receivers is defined as

$$\phi(\bar{\mathbf{m}}) = \max \left(\frac{1}{\Delta T} \int_{T_1}^{T_1 + \Delta T} |B(t, \bar{\mathbf{m}})|^2 dt \right), \quad 0 \leq T_1 \leq T - \Delta T, \quad (6)$$

where

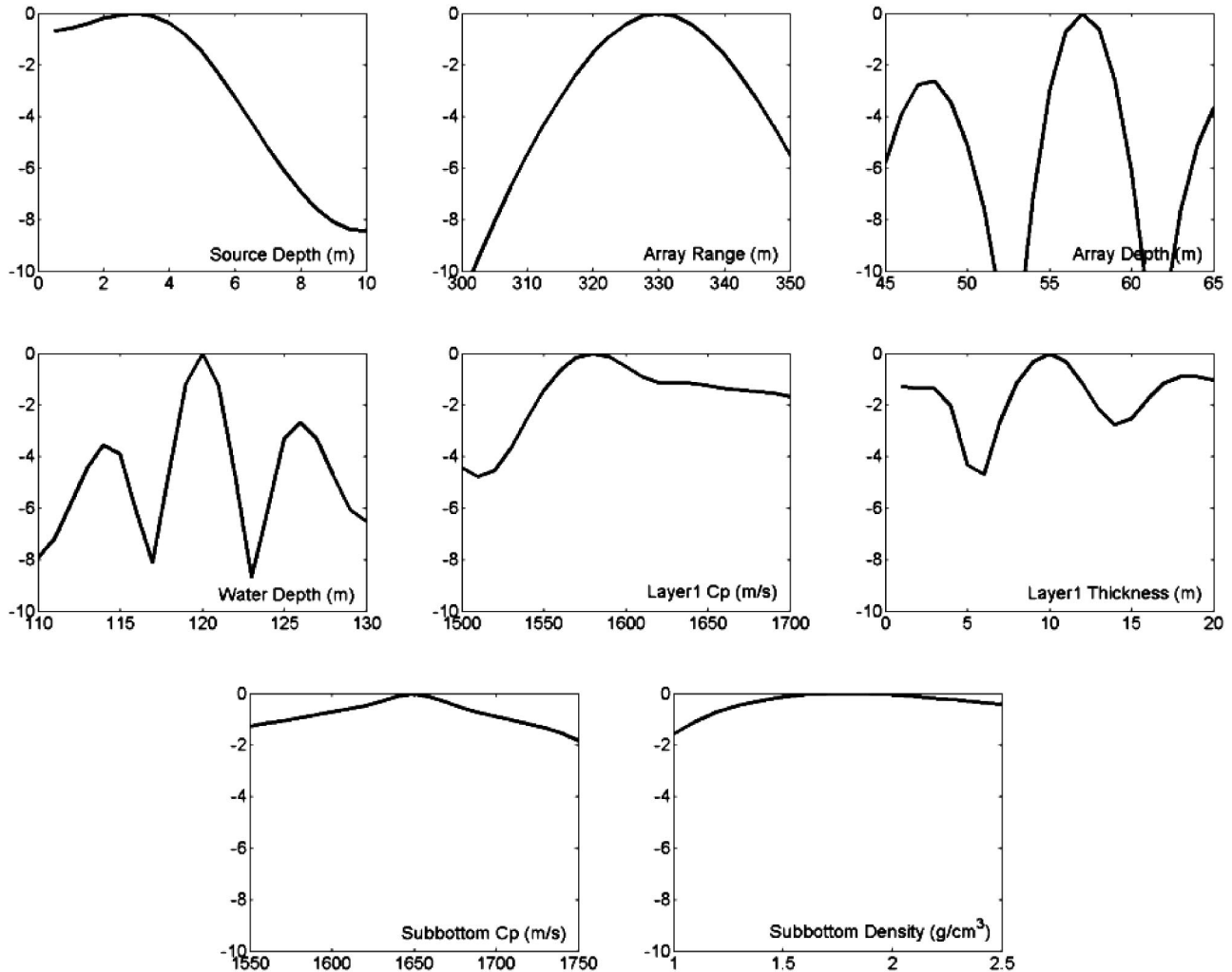


FIG. 4. Sensitivity curves of various parameters which show the objective function values evaluated at the parameter values given in the x -axis. Meanwhile, other parameters are fixed at their reference values given in Fig. 3. Note that the function values are given in dB normalized to the maximum value.

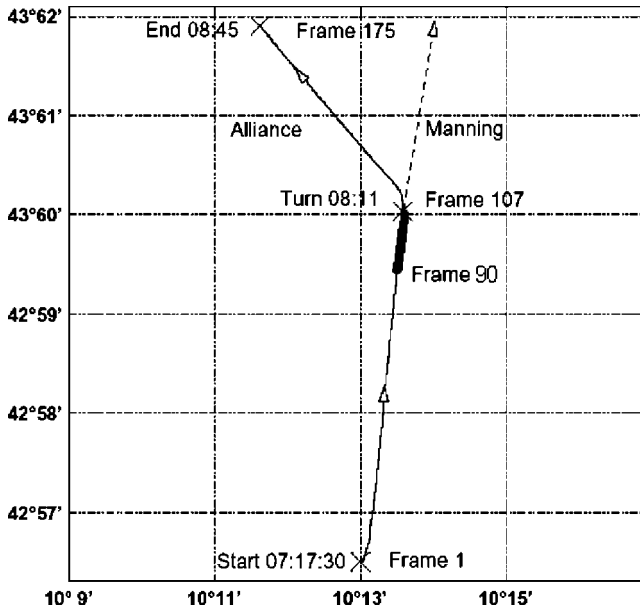


FIG. 5. The tracks of the RV *Alliance* and *Manning* during the experiment. The thick line indicates the track of *Manning* used as sources of opportunity in geoacoustic inversions.

$$B(t, \bar{\mathbf{m}}) = \frac{1}{N} \sum_{i=1}^N f_i(t, \bar{\mathbf{m}}). \quad (7)$$

A ray-based model is chosen for the forward model to calculate the objective function. In order to construct the impulse response in the ray-based model, we need information of amplitudes and phases of the rays that arrive at the receiver position within a predetermined error bound.⁶ It was found that a simple plane wave approximation is sufficient to calculate the phase and amplitude of a ray if the grazing angle of the ray is not near the critical angle, which corresponds to our inversion case presented in Sec. IV. However, if the grazing angle falls within near the critical angle, spherical reflection coefficients²⁴ instead of plane wave ones are calculated in the forward model in order to incorporate the spherical wave effects, especially for the low frequency component.

Since there are numerous results on matched field inversion with vertical line arrays, a relatively good understanding of the sensitivity exists. However, less information is known on the time-domain inversions using a towed system and ship noises. In order to get better insights on the objective function and its sensitivity characteristics with respect to the geoacoustic parameters, we investigated the behavior of the objective function as a function of the selected parameter with other parameters fixed at their reference values. This sensitivity test is commonly carried out in practice in geoacoustic inversions. Although the sensitivity estimates using one parameter may show biased results due to parameter couplings,²⁵ it is considered meaningful to compare the sensitivity for our time-domain inversion with that of other matched field inversions.

For the sensitivity test, we assumed that the tow ship was a sound source and emitted both the random noise and a 320 Hz tonal signal. The environment was the same as Fig. 3. The noise generated by a tow ship propagated through the

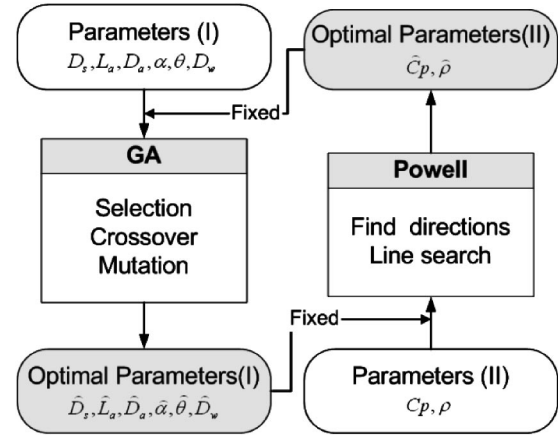


FIG. 6. Two-step optimization procedure.

water column and one layered ocean sediment and was measured on 32 receivers of the HLA. The power spectrums of measured signals for all receivers are also shown in Fig. 3.

Figure 4 shows the sensitivity curves of the parameters using 1 second data which correspond to 6000 time samples. The averaged time interval ΔT in Eq. (6) is chosen to be half the data length. The geometrical parameters such as the source position, the HLA configuration and the water depth show high sensitivities. The parameters of the layer and the sub-bottom show intermediate and low sensitivities, respectively. The sensitive parameters share a common characteristic in that they affect the travel times, i.e., phases, of the received signals. The least sensitive parameter which contributes to the wave amplitude is the density. The sound speed of the sub-bottom is also related with the amplitude in this simulation since a majority of the reflections from the sub-bottom interface occur before the total reflection range. In the simulated experimental setup, the covering angle of incidence with respect to vertical direction by the receivers ranges from 58° to 69° , but the critical angle of incidence is approximately 67° . The above results are mainly due to a nature of the matched impulse response filter. The physical meaning of the filtered output in Eq. (5) is the correlation between the measured signal and the simulated impulse response. The most influential factor in the correlation of the signals is the phase of each of the signal components. The differences in the phases result in a considerable reduction of correlation, but those of the amplitudes take little effect. Battle *et al.*⁵ calculated the sensitivity of the sound speed of half space with respect to the distance between the source and HLA. According to them, the sensitivity showed a rapid increase as the source–receiver range approaches the critical range. After the critical range where the dependence of the reflection coefficient switches from amplitude to phase, the sensitivity saturated and then oscillated showing higher values than before the critical range. This implies that the parameter affecting the phase of the signal is very sensitive in the inversion.

IV. EXPERIMENT AND PARAMETRIZATION

The MAPEX 2000 ship noise experiment was conducted by the SACLANT Undersea Research Center near the island

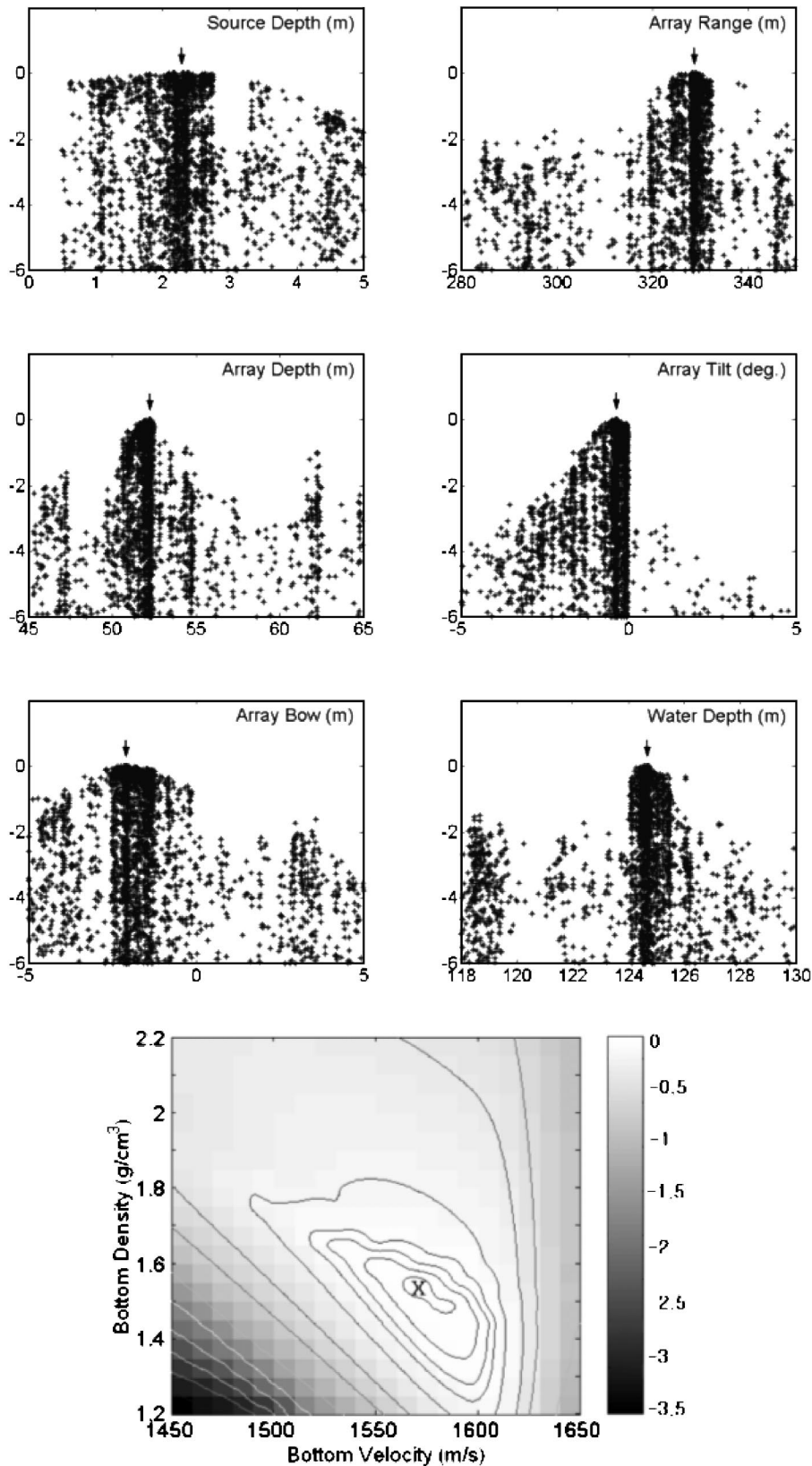


FIG. 7. Inversion results for data measured at 08:07:30 UTC (frame 100). The upper 6 figures are scatter plots of GA search and the lower two dimensional contour shows the objective function values which are re-evaluated based on the best parameters of GA and bottom parameters given in both x - and y -axes. Arrows and crosses represent the best solutions and all function values are given in dB normalized to the maximum value.

of Elba on November 29th. The data were received by a 254 m long HLA for 10 seconds every 30 seconds corresponding to data of 175 frames. The HLA was composed of 128 receivers with 2 m separations and was towed by the R/V *Alliance* along the track shown in Fig. 5.

From this figure, we can see that another R/V *Manning* was following the *Alliance* approximately 900 m behind along the same track from about frame 60 to frame 107. At frame 107, the *Alliance* started making a 45° turn, while the *Manning* proceeded on track as shown in Fig. 5. After frame

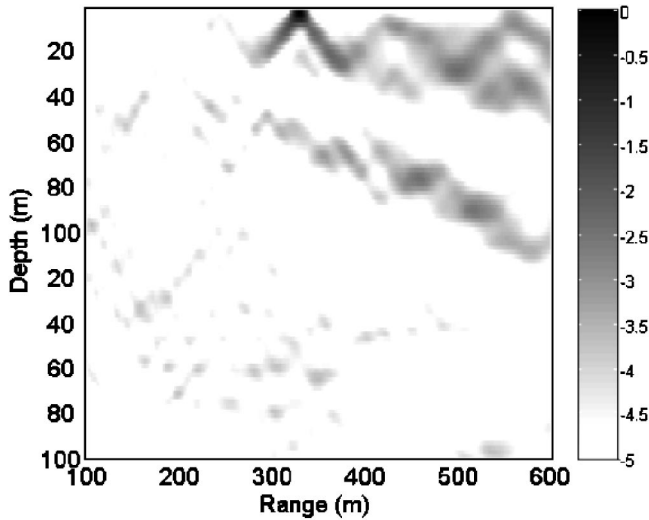


FIG. 8. Ambiguity surface of the source localization for the data measured at 08:07:30 UTC (frame 100) calculated using other inverted parameters except for the source location. The function values are given in dB normalized to the maximum value.

122, the *Manning* increased her speed and left the station. Since the distance between the *Alliance* and the nearest receivers to her was approximately 330 m, the horizontal range of the *Manning* from the tail of the array can be estimated to be about 320 m. Although the proximity of two ships may cause some interference, the *Manning* can still be utilized as a source of opportunity.

During the experiment, the depth-sounder recorded the bathymetry showing a slow growth from 116 m to 124 m for the entire track. The sound speed of the water column was kept almost constant at 1520 m/s according to expendable bathy-thermograph casts. The experimental sites have been studied for a long time and the bottom properties are well known. The bottom is comprised of a flat thin layer and a sub-bottom. The thin layer is composed of clay and clay-sand sediments and its thickness is approximately 2.5 m which is only a fraction of a wavelength at frequencies of a few hundred Hz. According to Gingras *et al.*,²⁶ the p-wave speed of the layer was estimated to increase linearly from 1505 m/s to 1556 m/s, which was *a posteriori* mean values based on the 40 observations. In each observation, however, the estimated upper and lower p-wave speeds of the layer showed high fluctuation. The estimated mean p-wave sound speed and density of the sub-bottom were 1578 m/s and 1.6 g/cm³, respectively. Therefore, the thin layer is disregarded in the present inversion and the environmental model of the inversion is treated as range-independent infinite half space. The effect of the thin layer, which is assumed to be small and insensitive to wave propagation, will be included in the infinite half bottom properties in a depth-averaged manner.

Even though the towed HLA is devised to be neutrally buoyant in the water, it still remains to be susceptible to deformations in shape during operations. Therefore, the array shape needs to be included in the inversion parameters for reliable inversion results. Battle *et al.*⁵ implemented the array shape deformation by defining the receiver depth as

$$z_i = \alpha \left(1 - \left(\frac{2x_i}{L} \right)^2 \right) - x_i \sin \theta, \quad (8)$$

where x_i and z_i are the relative positions of the receiver with respect to the center of the array whose length is L . The inversion parameters α and θ in Eq. (8) represent the array bow and tilt, respectively. The same definition is adopted for our inversion. Consequently, a total of eight parameters will be inverted in subsequent inversions; source depth D_s , distance between the source and the nearest receiver L_a , depth of the array center D_a , array tilt θ , array bow α , water depth D_w , bottom sound speed Cp , and bottom density ρ . The attenuation of bottom will be excluded in the inversion due to its extremely low sensitivity.

V. INVERSION USING SOURCE OF OPPORTUNITY

There were two available sources, *Alliance* and *Manning*, used in this experiment. Among these two, the *Manning* data measured from frame 90 to frame 107 which corresponds to the track of thick line in Fig. 5 were considered in this paper. The data of each frame consisted of 10 second signals. Among the 128 receivers, we used every fourth receiver data (32 receivers) for the inversion. In addition, we divided each signal into 10 segments of 1 second long and inverted all of them.

Frequent traffic was reported near the experimental site and the data are possibly contaminated by other ships. Therefore, it was recommended to reduce the unwanted interferences using a wavenumber filter before inversion.⁵ The wavenumber filter was implemented by a 2-D FFT for discrete time signals in a line array. It is noted when the wavenumber filtered data are used in the objective function (6), the simulated impulse response should also include only the corresponding rays. The wavenumber cutoff was set to include arrival angles falling within 50° of endfire to capture most of the expected multiple arrivals from the *Manning*.

The inversion was performed via an optimization process which searches the parameters to maximize the objective function in Eq. (6). Direct and surface reflected signals usually have large amplitudes. But they contain only the geometric information, such as the source and receiver positions, and thus are not useful for the inversion of bottom properties since they do not hit the bottom. If the geometry and bottom properties are inverted simultaneously, geometrical parameters will dominate over bottom parameters in the search process. In order to prevent such dominance in the optimization, the parameters should be inverted separately based on their sensitivities.

In geoacoustic inversion, global optimization methods such as the genetic algorithm and the simulated annealing are preferred due to their potential in finding a global optimum. However, they may have a drawback of slow convergence to the optimum. On the contrary, efficient local methods show quick and sure convergence to the optimum closest to the starting point, but it cannot guarantee to be the global optimum in the multi-modal function space. Presuming with reference to Fig. 4 that the multi-modality could be alleviated if we construct the objective function with only two parameters

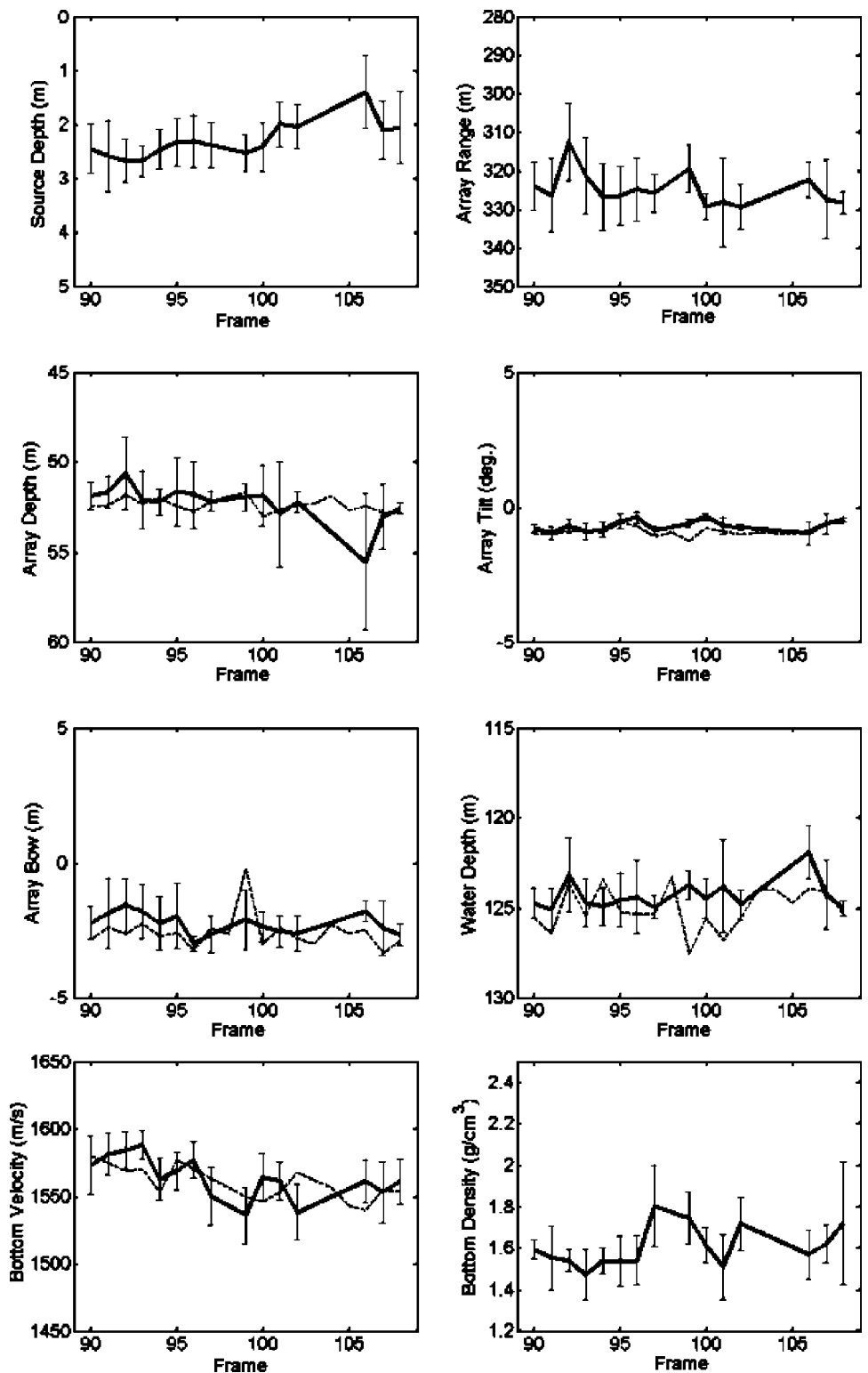


FIG. 9. The summary of inversion results which shows the mean and the standard deviation of 10 consecutive inversions of each frame by a thick line and error bars, respectively. The dashed lines represent reference values which were estimated by Battle *et al.* (Ref. 5) via a frequency domain matched field processing using tonal frequencies of 131.8, 262.9 and 306.9 Hz.

of bottom sound speed and bottom density, we took advantage of both the global and the local methods as follows.

At first, the genetic algorithm was applied to the sensitive parameters; the source depth, the array configurations and the water depth, and others fixed at certain values. After one generation of GA was constructed through the selection-crossover-mutation process, the remaining parameters were optimized using a local search method while already optimized parameters were kept at the best values among the

population. And the GA search for the next generation was performed again with the insensitive parameters fixed at optimal values of the local search. The above two-step search process was repeated until a pre-determined number of generations. This optimization procedure is summarized in Fig. 6.

In the inversion, the GA searches were performed for 50 generations and each generation contained 60 populations. The crossover and mutation occurred with probabilities of

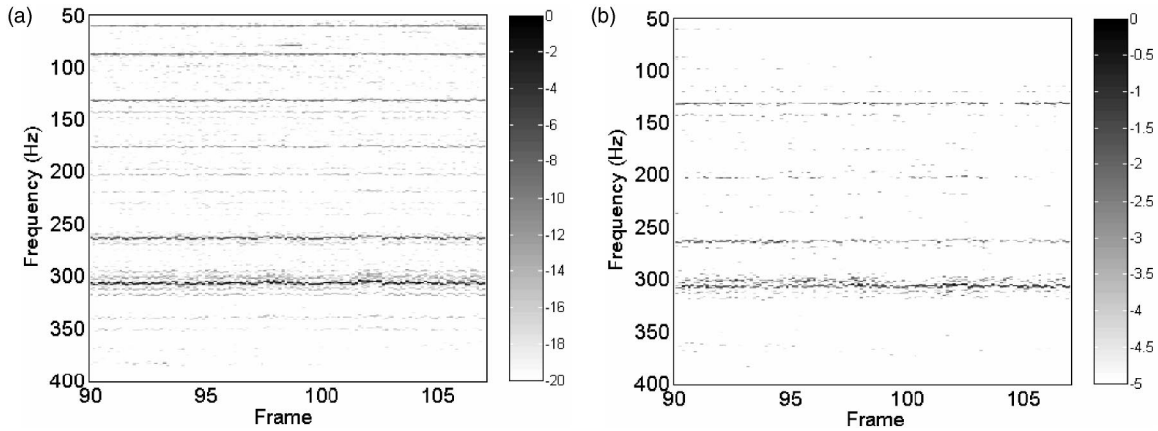


FIG. 10. A comparison of the spectrogram beamformed (a) and the Bartlett powers (b). The spectrogram was computed from beamformed time series in the *Manning* direction and the frequency dependent Bartlett powers were calculated with the inverted parameters of corresponding frame via time domain inversions. The function values are given in dB normalized to the maximum value.

0.9 and 0.05, respectively. Powell's method²⁷ was used as the local optimizer, since it does not need gradient information for finding its search direction.

A set of inversion results for a segment signal of frame 100 is presented in Fig. 7. Although one dimensional marginal *a posteriori* probability display of the results is a common practice,¹⁴ we plotted all objective function values with respect to the corresponding parameter values visited during the GA search. By doing so, we could obtain information about the optimizer (GA) such as its behavior related to the parameter sensitivity and sampled parameter space.¹⁰ For the insensitive bottom parameters, which are the main targets of the inversion, we re-evaluated the objective functions based on the best results of GA and showed them in a two dimensional contour plot in Fig. 7. In this figure, the arrows and the cross indicate the best solutions that were found during the inversion procedure. In the scatter plots of Fig. 7, all of the parameters of GA seem to have been resolved well. Besides, the two dimensional contour which shows the mono-modality of the objective function space may be a result that supports that the local optimization is useful.

For an additional check of the validity for the inverted parameter values, we simply performed a back-propagation of the received signals from the array using these environmental parameter results of Fig. 7. The ambiguity surface in Fig. 8 shows a good focus of the back-propagated signals to the correct source position, considering that the draft of *Manning* and estimated range from the nearest receiver to her are 2.5 m and 320 m, respectively.

In Fig. 9, we summarize the inversion results of all frames by the statistical mean and standard deviation estimated from 10 multiple inversions at each frame. The ranges of the y-axis are the same with the search intervals in the optimization procedures. The standard deviations imply uncertainties of the inversion which may arise from the varying data SNR and environments during the experiment and/or from the inherent errors of the optimizer that may occur if the search space is vast.²⁸ However, the mean estimates seem to be consistent especially for geometric parameters irrespective of their uncertainties at each frame. The relatively high variability in the bottom densities seem to be due to its in-

sensitivity to the wave propagation rather than to range dependencies, considering the experimental site has been known to be range independent. Besides, good agreements of our estimates with the reference values represented by the dashed lines in the figure are found. The reference values were estimated by Battle *et al.*⁵ via a frequency domain matched field processing using tonal frequencies of 131.8, 262.9 and 306.9 Hz, which were identified as prominent peaks in the spectrogram of Fig. 10(a). The spectrogram was computed from beamformed time series in the *Manning* direction.

In this paper, we define the Bartlett power as a function of frequency f by

$$B_i(f) = \frac{\mathbf{w}^\dagger(f, \mathbf{m}_i) \mathbf{R}_i(f) \mathbf{w}(f, \mathbf{m}_i)}{\text{tr}[\mathbf{R}_i(f)] \|\mathbf{w}(f, \mathbf{m}_i)\|^2}, \quad (9)$$

where \mathbf{R}_i is the estimated cross spectral density matrix of the i th segmented signals and $\mathbf{w}(\mathbf{m}_i)$ represents the replica vector calculated with the inverted parameters of the corresponding segment. Figure 10(b) shows the calculated Bartlett power. Comparing it with the beamformed spectrogram of Fig. 10(a), we can see a similarity between them. From the frequency dependent behavior of the Bartlett power, it seems that a few frequencies played dominant roles in the inversion and they coincide with those used by Battle *et al.*, which naturally explains the agreements between the two different inversion results. Therefore, we can conclude that the proposed time domain inversion takes advantage of dominant frequencies of the source signature automatically.

VI. CONCLUSION

Predicting bottom properties is important for various acoustic applications in shallow waters due to the significance of the bottom interaction. In this paper we propose a feasible geoacoustic inversion method using a towed HLA and ship noises as a source of opportunity. Such a simple system may make immediate use of geoacoustic parameters as an input to various sonar operations possible.

The inversion scheme is summarized as follows. The objective function was defined using a matched impulse re-

sponse filter for the unknown source signal. The filtered output was implemented by back-propagating the received and time-reversed signals using a ray-based forward model. Then, a series of global and local searches was performed using GA and Powell's method, respectively, to find the optimum parameters.

The present method was applied to the experimental data of the MAPEX2000 noise experiment. The *Manning* data were inverted successively and their results showed good consistency for all parameters. The inverted parameters were compared with the reference values estimated via a frequency domain matched field processing using three dominant tonal frequencies and results showed good agreements. In addition, analyses of Bartlett power revealed that the proposed time domain inversion takes advantage of dominant frequencies of the source signature automatically.

ACKNOWLEDGEMENT

This work was supported by the Basic Research Program of KORDI/KRISO.

- ¹N. R. Chapman and A. Tolstoy, "Special issue: Benchmarking geoaacoustic inversion methods," *J. Comput. Acoust.* **6**, 1–289 (1998).
- ²N. R. Chapman and M. Taroudakis, "Special issue: Geoaacoustic inversion in shallow water," *J. Comput. Acoust.* **8**, 259–388 (2000).
- ³W. A. Kuperman, M. F. Werby, K. E. Gilbert, and G. J. Tango, "Beam forming on bottom-interacting tow-ship noise," *IEEE J. Ocean. Eng.* **10**, 290–298 (1985).
- ⁴M. Siderius, P. Nielsen, and P. Gerstoft, "Range-dependent seabed characterization by inversion of acoustic data from a towed receiver array," *J. Acoust. Soc. Am.* **112**, 1523–1535 (2002).
- ⁵D. Battle, P. Gerstoft, W. A. Kuperman, W. S. Hodgkiss, and M. Siderius, "Geoacoustic inversion of tow-ship noise using matched-field processing," *IEEE J. Ocean. Eng.* **28**, 454–467 (2003).
- ⁶C. Park, W. Seong, P. Gerstoft, and M. Siderius, "Time domain geoaacoustic inversion of high-frequency chirp signal from a simple towed system," *IEEE J. Ocean. Eng.* **28**, 468–478 (2003).
- ⁷N. R. Chapman, R. M. Dizaji, and R. L. Kirilin, "Geoacoustic inversion using broad band ship noise," *Proceedings of the 5Pth ECUA*, Lyon, 2000, pp. 787–792.
- ⁸D. P. Knobles and R. A. Koch, "A time series analysis of sound propagation in a strongly multipath shallow water environment with an adiabatic normal mode approach," *IEEE J. Ocean. Eng.* **21**, 290–298 (1985).
- ⁹J.-P. Hermand, "Broad-band geoaacoustic inversion in shallow water from waveguide impulse response measurements on a single hydrophone: theory and experimental results," *IEEE J. Ocean. Eng.* **24**, 41–66 (1999).
- ¹⁰L. Jaschke and N. R. Chapman, "Matched field inversion of broadband data using the freeze bath method," *J. Acoust. Soc. Am.* **106**, 1838–1851 (1999).
- ¹¹C. W. Holland and J. Osler, "High-resolution geoaacoustic inversion in shallow water: a joint time- and frequency-domain technique," *J. Acoust. Soc. Am.* **107**, 1263–1279 (2000).
- ¹²P. Pignot and N. R. Chapman, "Tomographic inversion of geoaacoustic properties in a range-dependent shallow-water environment," *J. Acoust. Soc. Am.* **110**, 1338–1348 (2001).
- ¹³N. R. Chapman, J. Desert, A. Agarwal, Y. Stephan, and X. Demoulin, "Estimation of seabed models by inversion of broadband acoustic data," *Proceedings of the 6Pth ECUA*, Gdansk, 2002, pp. 477–481.
- ¹⁴P. Gerstoft, "Inversion of seismoacoustic data using genetic algorithms and a *a posteriori* probability distribution," *J. Acoust. Soc. Am.* **95**, 770–782 (1994).
- ¹⁵S. E. Dosso, M. L. Yeremey, J. M. Ozard, and N. R. Chapman, "Estimation of ocean-bottom properties by matched-field inversion of acoustic field data," *IEEE J. Ocean. Eng.* **18**, 232–239 (1993).
- ¹⁶P. Gerstoft, "Inversion of acoustic data using a combination of genetic algorithm and the Gauss-Newton approach," *J. Acoust. Soc. Am.* **97**, 2181–2190 (1995).
- ¹⁷M. R. Fallat and S. E. Dosso, "Geoacoustic inversion via local, global, and hybrid algorithms," *J. Acoust. Soc. Am.* **105**, 3219–3230 (1999).
- ¹⁸T. B. Nielsen, "An iterative implementation of rotated coordinates for inverse problems," *J. Acoust. Soc. Am.* **113**, 2574–2586 (2003).
- ¹⁹A. Parvulescu and C. S. Clay, "Reproducibility of signal transmission in the ocean," *Radio Electron. Eng.* **29**, 223–228 (1965).
- ²⁰C. S. Clay, "Optimum time domain signal transmission and source location in a waveguide," *J. Acoust. Soc. Am.* **81**, 660–664 (1987).
- ²¹R. K. Brienzo and W. S. Hodgkiss, "Broadband matched-field processing," *J. Acoust. Soc. Am.* **94**, 2821–2831 (1993).
- ²²D. R. Jackson and D. R. Dowling, "Phase conjugation in underwater acoustics," *J. Acoust. Soc. Am.* **89**, 171–181 (1991).
- ²³D. R. Dowling, "Acoustic pulse compression using passive phase-conjugate processing," *J. Acoust. Soc. Am.* **95**, 1450–1458 (1994).
- ²⁴L. M. Brekhovskikh, *Waves in Layered Media*, 2nd ed. (Academic, New York, 1980).
- ²⁵S. E. Dosso, "Quantifying uncertainty in geoaacoustic inversion. I. A fast Gibbs sampler approach," *J. Acoust. Soc. Am.* **111**, 129–142 (2002).
- ²⁶D. F. Gingras and P. Gerstoft, "Inversion for geometric and geoaacoustic parameters in shallow water: Experimental results," *J. Acoust. Soc. Am.* **97**, 3589–3598 (1995).
- ²⁷W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. (Cambridge University Press, Cambridge, 1992).
- ²⁸M. Snellen, D. G. Simons, M. Siderius, J. Sellschopp, and P. L. Nielsen, "An evaluation of the accuracy of shallow water matched field inversion results," *J. Acoust. Soc. Am.* **109**, 514–527 (2001).

Effects of environmental uncertainties on sonar detection performance prediction

Liewei Sha and Loren W. Nolte^{a)}

Department of Electrical and Computer Engineering, Box 90291, Duke University, Durham, North Carolina 27708-0291

(Received 8 June 2003; accepted for publication 31 January 2005)

The development of effective passive sonar systems depends upon the ability to accurately predict the performance of sonar detection algorithms in realistic ocean environments. Such environments are typically characterized by a high degree of uncertainty, thus limiting the usefulness of performance prediction approaches that assume a deterministic environment. Here we derive closed-form receiver operating characteristic (ROC) expressions for an optimal Bayesian detector and for several typical suboptimal detectors, based on a statistical model of environmental uncertainty. Various scenarios extended from an NRL benchmark shallow-water model were used to check the analytical ROC expressions and to illustrate the effect of environmental uncertainty on detection performance. The results showed that (1) optimal detection performance in an uncertain environment in diffuse noise depends primarily on the signal-to-noise ratio at the receivers and the rank of the signal matrix, where the rank is an effective representation of the scale of environmental uncertainty; (2) the ROC expression for the optimal Bayesian detector provides a more realistic performance upper bound than that obtained from conventional sonar equations that do not incorporate environmental uncertainty; and (3) detection performance predictions can be performed much faster than with commonly used numerical methods such as Monte Carlo performance evaluations. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1875653]

PACS numbers: 43.30.Wi, 43.60.Gk, 43.60.Uv [WLS]

Pages: 1942–1953

I. INTRODUCTION

To detect the presence of an underwater acoustic source, an optimal detector must combine array data with knowledge about the source position, the propagation channel, and the receiving array configuration. Although the receiving array configuration is typically known, the knowledge of sensor positions is imperfect. The source position is usually unknown. Of particular concern is the uncertainty associated with the propagation channel. The spatial and temporal variations in the channel parameters, such as the ocean bottom depth and the sound speed profile, make accurate characterization of the channel difficult. The set of propagation channel parameters together with the source position parameters constitute an uncertain ocean acoustic environment. Environmental uncertainty can cause severe performance degradation to sonar detection algorithms that rely on precise knowledge of the environmental parameters.^{1–4} Sonar detection performance prediction methods that fail to consider environmental uncertainty do not capture the performance degradation it causes, leading to overestimation of detection performance. Thus, a practical sonar detection performance prediction method that incorporates the effects of environmental uncertainty, and consequently predicts sonar detection performance more accurately, is of practical interest.

Detection performance prediction typically involves both a detection algorithm and a methodology for evaluating the algorithm's performance. Several algorithms that are ro-

bust to environmental uncertainty have been proposed, including the spatial stationarity test,⁵ the focalization approach,⁶ and generalized matched field processing (MFP) algorithms, etc. In the focalization approach,⁶ key environmental parameters are adjusted in an attempt to bring the source into focus before localization. Generalized MFP algorithms include the minimum variance beamformer with multiple neighboring location constraints (MV_NCL),³ the minimum variance beamformer with environmental perturbation constraints (MV_EPC),⁷ the optimum uncertain field processor (OUFP),^{8,9} and the reduced-rank MFP algorithms.¹⁰ Although most of the MFP algorithms were originally proposed for source localization, with a few modifications they can be used to detect the presence of a target. The performance of detection algorithms for an uncertain ocean environment has been evaluated in the past. However, few statistically valid detection performance predictions have been reported. The robustness of the spatial stationarity test⁵ was demonstrated using Mediterranean vertical array data.¹¹ The detection performance of reduced-rank MFP algorithms was tested using data from the Santa Barbara Channel experiment.¹⁰ These performance analyses generate *ad hoc* assessments but not general conclusions about the effects of environmental uncertainty on the detection performance. Some performance evaluations have used a quantitative environmental mismatch model.^{4,2,12} In one approach,⁴ the degradation of the localization performance of a matched-mode-processor was determined as a function of increasing bottom depth mismatch. Their predictions can be improved by including environmental uncertainties other than ocean depth mismatch or sound speed profile mismatch. Finally, most detection perfor-

^{a)} Author to whom correspondence should be addressed. Electronic mail: lwn@ee.duke.edu

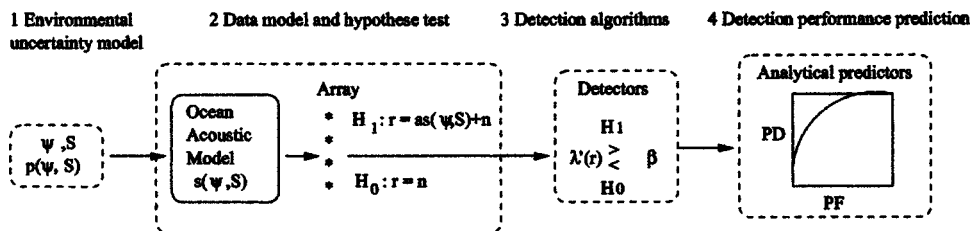


FIG. 1. Detection performance prediction flow chart.

mance evaluations have been implemented using Monte Carlo simulations.^{9,13} In using the Monte Carlo approach to obtain the detection performance, one needs to generate Monte Carlo data samples over environmental uncertainties as well as implement Monte Carlo integration over the uncertainties of the data, which can be very computationally intensive. Furthermore, the Monte Carlo approach provides no direct insight into fundamental parameters that impact sonar detection performance.

Here we derive analytical performance predictions for an optimal Bayesian detector^{8,9} and for several suboptimal detectors, based on a statistical model of environmental uncertainty. The performance metric used in our prediction methodology is the receiver operating characteristic (ROC). The ROC is more relevant to source detection than metrics such as array gain. The optimal Bayesian detector is a modified version of the OUPF,^{8,9,14–16} which fully incorporates the *a priori* knowledge of the environment. Because the optimal Bayesian detector considers the presence of environmental uncertainty and fully incorporates the *a priori* knowledge, its performance prediction provides a practical performance upper bound. The approximate closed-form ROC expression presented in this paper for the optimal Bayesian detector, referred to as the optimal Bayesian predictor, is a new contribution. The optimal Bayesian predictor reveals that optimal detection performance in an uncertain environment in the presence of diffuse noise depends primarily on the signal-to-noise ratio (SNR) and on the approximate rank of the signal matrix, where the rank reflects the scale of environmental uncertainty. The computation is much faster than using Monte Carlo simulations. Another contribution of this paper is the derivation of precise analytical ROC expressions for the matched-ocean detector, the mean-ocean detector, and the energy detector as a function of environmental uncertainties. The analytical ROC expression for the mean-ocean detector can also be applied generally to predict the detection performance of a beamformer, when the signal wavefront, i.e., the normalized acoustic field, is mismatched with the data.

The organization of the paper is as follows. The detection performance prediction methodology is presented in Sec. II. The optimal Bayesian detector and several suboptimal detectors are presented in Sec. III. The analytical ROC expressions for those detection algorithms are derived in Sec. IV. In Sec. V, a group of simulated uncertain shallow-water environments are used to check the analytical detection performance predictions and to analyze the effects of environmental uncertainty. These simulated environments are extensions from a general mismatch benchmark environment

(“genlmis”) proposed in the May 1993 NRL Workshop on Acoustic Models in Signal Processing.^{5,9,17}

II. DETECTION PERFORMANCE PREDICTION METHODOLOGY

The detection performance prediction methodology is illustrated in Fig. 1, where the symbols are listed in Table I. Environmental uncertainty is explicitly modeled using a statistical approach. An ocean acoustic propagation model is used to transform environmental uncertainty into the uncertainty of the received array signal. The sonar detection problem is placed in a statistical hypothesis test framework. The detection algorithms under investigations are all in the form of a test statistic that is compared with a threshold to determine which hypothesis is true. Analytical detection performance predictions are derived using the ROC metric. This section presents the basic part of the methodology, which includes the environmental uncertainty model, the hypothesis test model, and the performance metric.

A. Environmental uncertainty model

The purpose of modeling environmental uncertainty is to quantitatively analyze its effects on detection performance. Here environmental uncertainty is quantified by assigning an *a priori* probability density function to both the uncertain channel parameters and the unknown source position parameters, over a finite *range* of possible values. The symbol Ψ is used to represent the channel parameters and the symbol \mathbf{S} is used to represent the source position parameters. It is assumed that Ψ and \mathbf{S} are statistically independent *a priori* with uniform probability density functions denoted by $p(\Psi)$ and $p(\mathbf{S})$, respectively. The advantage of this model com-

TABLE I. Symbol definitions for detection performance prediction methodology.

Symbol	Definition
Ψ	Acoustic propagation channel parameters
\mathbf{S}	Source position parameters
\mathbf{r}	Received frequency domain data
$\mathbf{H}(\Psi, \mathbf{S})$	Replica field
a	Frequency domain signal complex amplitude
σ_a^2	Source amplitude variance
\mathbf{n}	Frequency domain diffuse noise
σ_n^2	Noise variance
$p(\Psi, \mathbf{S})$	Probability density function of the environmental parameters
$\lambda'(\mathbf{r})$	Test statistic
β	Detection threshold
P_D	Probability of detection
P_F	Probability of false alarm

pared to the mismatched environmental uncertainty model^{4,2,12} is that it enables the performance prediction algorithm to incorporate an ensemble of possible realizations of the uncertain environment.

B. Hypothesis test

The detection problem is formulated as a decision problem: given the measurement data \mathbf{r} , accept the “noise only” hypothesis H_0 , or reject it to accept the “signal present” hypothesis H_1 , in the presence of channel parameter uncertainties and source position uncertainty. This differs from the source localization problem that often assumes the H_1 hypothesis only.

The two hypotheses are defined as

$$\begin{aligned} H_1: \mathbf{r} &= a_0 \mathbf{H}(\Psi, \mathbf{S}) + \mathbf{n}_0, \\ H_0: \mathbf{r} &= \mathbf{n}_0, \\ \mathbf{n}_0 &\sim \mathcal{N}(0, 2\sigma_n^2 \mathbf{I}_N), \quad a_0 \sim \mathcal{N}(0, 2\sigma_a^2). \end{aligned} \quad (1)$$

The data \mathbf{r} and the diffuse noise \mathbf{n}_0 are both $N \times 1$ spatial column vectors, corresponding to N spatially distributed sensors in the receiving array. The i th component of the data vector \mathbf{r} is assumed to be a narrow-band Fourier transform of the snapshot received by the sensor at the i th location. The ocean acoustic transfer function,¹⁶ $\mathbf{H}(\Psi, \mathbf{S})$, is the complex-valued acoustic pressure field sampled at the receiving array of N sensors in the frequency domain, given ocean environmental parameters Ψ and source position parameters \mathbf{S} . Note that $\mathbf{H}(\Psi, \mathbf{S})$ is known in the applied physics literature as the Green’s function, and is frequently denoted by \mathbf{G} . The scalar a_0 is a circulant complex Gaussian random variable that represents the unknown frequency-domain source amplitude and phase, with zero mean and known variance $2\sigma_a^2$. The diffuse noise present at the receiving array is represented by \mathbf{n}_0 in the frequency domain, which is a circulant complex Gaussian random vector. Denoting the expectation by $\mathbf{E}(\cdot)$ and the covariance matrix by $\mathbf{Cov}(\cdot)$, the assumptions are $\mathbf{E}(\mathbf{n}_0) = \mathbf{0}$ and $\mathbf{Cov}(\mathbf{n}_0) = 2\sigma_n^2 \mathbf{I}_N$, where \mathbf{I}_N is an identity matrix of size N .

The acoustic pressure field generated by a point source at depth z_s and range r_s is computed using normal mode theory,¹⁸

$$\begin{aligned} p(z; r_s, z_s) &= p_0 \sum_{m=1}^K \Phi_m(z_s) \Phi_m(z) \\ &\quad \times \exp(-jk_m r_s) / \sqrt{k_m r_s}, \end{aligned} \quad (2)$$

where z is the depth of the receiver and p_0 is a normalization constant. The number of propagating acoustic modes is represented by K . The terms k_m and Φ_m are the m th eigenvalue and eigenfunction of the Sturm-Liouville problem, which can be calculated using Porter’s KRAKEN¹⁹ code.

In addition, the signal wavefront, $\mathbf{s}(\Psi, \mathbf{S})$, i.e., the normalized acoustic field, is given by

$$\mathbf{s}(\Psi, \mathbf{S}) = \mathbf{H}(\Psi, \mathbf{S}) / \|\mathbf{H}(\Psi, \mathbf{S})\|. \quad (3)$$

Here $\|\cdot\|$ means the two-norm on a complex vector space.

The signal-to-noise ratio at the receivers is defined as

$$\text{SNR}(\Psi, \mathbf{S}) = \frac{\sigma_a^2 \|\mathbf{H}(\Psi, \mathbf{S})\|^2}{\sigma_n^2}. \quad (4)$$

The SNR includes the factors of the source variance, the diffuse noise variance, and the array gain. It is not an element-level SNR but an array-level SNR.

C. Performance metric

The detection performance metric, ROC, is a plot of the probability of detection, P_D , as a function of the probability of false alarm, P_F . From signal detection theory, P_D and P_F are defined as

$$\begin{aligned} P_F(\beta) &= P(\lambda'(\mathbf{r}|H_0) \geq \beta), \\ P_D(\beta) &= P(\lambda'(\mathbf{r}|H_1) \geq \beta), \end{aligned} \quad (5)$$

where $\lambda'(\cdot)$ is the test statistic produced by a detection algorithm and β is the threshold which is determined by the decision criterion. The decision criterion is beyond the scope of this paper. Only the test statistics and the ROC performance are discussed.

When the probability density functions $p(\lambda'|H_0)$ and $p(\lambda'|H_1)$ are available, P_F and P_D can be computed using

$$\begin{aligned} P_F(\beta) &= \int_{\beta}^{\infty} d\lambda' p(\lambda'|H_0), \\ P_D(\beta) &= \int_{\beta}^{\infty} d\lambda' p(\lambda'|H_1). \end{aligned} \quad (6)$$

III. DETECTION ALGORITHMS

Detection algorithms can be categorized according to the level of utilization of the *a priori* knowledge. In this section, four detection algorithms are presented: the optimal Bayesian detector, the matched-ocean detector, the mean-ocean detector, and an energy detector. The matched-ocean detector is not a realistic algorithm, because a complete match between the assumed signal wavefront and the data is practically impossible. We include it to provide a detection performance prediction upper bound for an ideal situation. The mean-ocean detector is a modified version of the conventional, or Bartlett, MFP processor.^{18,20} It represents the algorithms that might use mismatched model parameters. The energy detector represents the algorithms that do not utilize the *a priori* knowledge of the environment at all.

A. Optimal Bayesian detector

The optimal Bayesian detector differs from the OUPF^{8,9} in that it implements the likelihood ratio, while the OUPF forms the maximum *a posteriori* probability (MAP) intended for localization. The derivation of the likelihood ratio is presented here.

Under the H_0 hypothesis, the probability density function of the data at the receiving array is assumed to be

$$p(\mathbf{r}|H_0) = \frac{1}{(2\pi)^N \sigma_n^{2N}} \exp\left(-\frac{\mathbf{r}^\dagger \mathbf{r}}{2\sigma_n^2}\right). \quad (7)$$

Under the H_1 hypothesis, the probability density function of the data conditional to the channel and source position parameters is $p(\mathbf{r}|H_1, \Psi, \mathbf{S})$. The data, conditional to the environmental parameters, is a complex Gaussian random vector. Its probability density function depends only on the mean vector and the covariance matrix, which are

$$\begin{aligned} \mathbf{E}(\mathbf{r}|H_1, \Psi, \mathbf{S}) &= \mathbf{0}, \\ \mathbf{C}_r &= \text{Cov}(\mathbf{r}|H_1, \Psi, \mathbf{S}) \\ &= 2\sigma_n^2 \|\mathbf{H}(\Psi, \mathbf{S})\|^2 \mathbf{s}(\Psi, \mathbf{S}) \mathbf{s}(\Psi, \mathbf{S})^\dagger + 2\sigma_n^2 \mathbf{I}_N. \end{aligned} \quad (8)$$

Applying Woodbury's identity²¹ and the definition of the SNR [Eq. (4)],

$$\mathbf{C}_r^{-1} = (2\sigma_n^2)^{-1} \left(\mathbf{I} - \frac{\text{SNR}(\Psi, \mathbf{S}) \mathbf{s}(\Psi, \mathbf{S}) \mathbf{s}(\Psi, \mathbf{S})^\dagger}{1 + \text{SNR}(\Psi, \mathbf{S})} \right). \quad (9)$$

The determinant of \mathbf{C}_r is the product of all its eigenvalues. Considering that \mathbf{C}_r has $N-1$ identical eigenvalues $2\sigma_n^2$ and one eigenvalue $2\sigma_n^2 \|\mathbf{H}(\Psi, \mathbf{S})\|^2 \mathbf{s}(\Psi, \mathbf{S})^\dagger \mathbf{s}(\Psi, \mathbf{S}) + 2\sigma_n^2$, and using the definition of the SNR,

$$|\mathbf{C}_r| = (2\sigma_n^2)^N (\text{SNR}(\Psi, \mathbf{S}) + 1). \quad (10)$$

Using Eqs. (9) and (10),

$$\begin{aligned} p(\mathbf{r}|H_1, \Psi, \mathbf{S}) &= \exp(-\mathbf{r}^\dagger \mathbf{C}_r^{-1} \mathbf{r}) |\pi \mathbf{C}_r|^{-1}, \\ &= \exp(-\mathbf{r}^\dagger \mathbf{r} / 2\sigma_n^2 + B(\Psi, \mathbf{S}) |\mathbf{r}^\dagger \mathbf{s}(\Psi, \mathbf{S})|^2) / A(\Psi, \mathbf{S}) \\ &\quad \times (2\pi\sigma_n^2)^N, \\ A(\Psi, \mathbf{S}) &= \text{SNR}(\Psi, \mathbf{S}) + 1, \\ B(\Psi, \mathbf{S}) &= \text{SNR}(\Psi, \mathbf{S}) / 2\sigma_n^2 (\text{SNR}(\Psi, \mathbf{S}) + 1). \end{aligned} \quad (11)$$

The conditional likelihood ratio is

$$\lambda(\mathbf{r}|\Psi, \mathbf{S}) = \frac{p(\mathbf{r}|H_1, \Psi, \mathbf{S})}{p(\mathbf{r}|H_0)}. \quad (12)$$

The likelihood ratio combines the probability density function of the environmental parameters with the conditional likelihood ratio. Assuming Ψ and \mathbf{S} are statistically independent *a priori*, the likelihood ratio becomes

$$\begin{aligned} \lambda(\mathbf{r}) &= \int_{\Omega_\Psi} \int_{\Omega_S} d\Psi d\mathbf{S} \lambda(\mathbf{r}|\Psi, \mathbf{S}) p(\Psi) p(\mathbf{S}), \\ \lambda(\mathbf{r}|\Psi, \mathbf{S}) &= \exp(B(\Psi, \mathbf{S}) |\mathbf{r}^\dagger \mathbf{s}(\Psi, \mathbf{S})|^2) / A(\Psi, \mathbf{S}), \end{aligned} \quad (13)$$

where Ω_Ψ and Ω_S represent the ranges of environmental uncertainties.

The optimal Bayesian detector is the likelihood ratio test

$$\lambda(\mathbf{r}) \underset{H_0}{\overset{H_1}{\geq}} \beta, \quad (14)$$

where β is the threshold.

B. Matched-ocean detector

When the ocean environmental parameters and the source position are known, the probability density functions

$p(\Psi)$ and $p(\mathbf{S})$ are delta functions. This is the limiting case of the optimal Bayesian detector, where there is no environmental uncertainty. Substituting delta functions for $p(\Psi)$ and $p(\mathbf{S})$ in Eq. (13) generates $\lambda(\mathbf{r}) = \lambda(\mathbf{r}|\Psi, \mathbf{S})$. Since $A(\Psi, \mathbf{S})$ and $B(\Psi, \mathbf{S})$ in Eq. (13) are now positive constants, $\lambda'(\mathbf{r})_{\text{matched}}$, a monotonic function of the likelihood ratio, is also optimum.

$$\lambda'(\mathbf{r})_{\text{matched}} = |\mathbf{r}^\dagger \mathbf{s}(\Psi_{\text{matched}}, \mathbf{S}_{\text{matched}})|^2. \quad (15)$$

The subscript ‘‘matched’’ is used to indicate that the environmental parameters in the model are matched with the real data. The performance prediction based on the matched-ocean detector provides an ideal performance upper bound. However, it can be too optimistic to use as a predictor of detection performance in a realistic uncertain ocean environment.

C. Mean-ocean detector

The mean-ocean detector uses a test statistic of the same form as the matched-ocean detector. It assumes that the detector knows only mean values of the environmental parameters, which are mismatched with the true environmental parameters. Subscript ‘‘mean’’ is used to indicate this assumption. The test statistic for the mean-ocean detector is expressed as

$$\lambda'(\mathbf{r})_{\text{mean}} = |\mathbf{r}^\dagger \mathbf{s}(\Psi_{\text{mean}}, \mathbf{S}_{\text{mean}})|^2. \quad (16)$$

D. Energy detector

The energy detection algorithm is a conventional approach to the detection of ocean acoustic sources. The energy detector does not include environmental uncertainties in its algorithm. Although robust to environmental uncertainties, it does not exploit these uncertainties. Using the data model defined in the frequency domain [Eq. (1)], the test statistic for the energy detector is

$$\lambda'(\mathbf{r})_{\text{energy}} = \mathbf{r}^\dagger \mathbf{r}. \quad (17)$$

IV. ANALYTICAL ROC PERFORMANCE PREDICTIONS

In this section, analytical ROC performance expressions are derived for each of the four detection algorithms, in the presence of various amounts of ocean environmental and source position uncertainties. These analytical ROC expressions are referred to as the optimal Bayesian predictor, the matched-ocean predictor, the mean-ocean predictor, and the energy predictor. The concept of the signal matrix is introduced first, since it is central to each of the sonar performance predictors.

A. Signal matrix

The signal wavefront, $\mathbf{s}(\Psi, \mathbf{S})$, is a function of the continuous random variables Ψ and \mathbf{S} . To computationally predict the performance of the detection algorithms in an uncertain environment, a discrete representation of the uncertain signal wavefront is needed. A matrix composed of M realizations of the signal wavefront due to uncertainties of the environmental parameters, referred to as the signal matrix \mathfrak{R} , is defined by applying Monte Carlo sampling techniques:

$$\begin{aligned} \mathfrak{X} &= [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_M], \\ &= [s((\Psi, \mathbf{S})_1), s((\Psi, \mathbf{S})_2), \dots, s((\Psi, \mathbf{S})_M)], \end{aligned} \quad (18)$$

where $(\Psi, \mathbf{S})_i$, $i = 1, \dots, M$, corresponds to the i th realization of the environmental parameters (Ψ, \mathbf{S}) , which follow the uniform probability density function $p(\Psi)p(\mathbf{S})$ that is defined on uncertainty ranges Ω_Ψ and Ω_S . Note that although the signal matrix is used to derive the ROC expressions, it is shown later that only its energy and rank information are needed in the resultant approximate ROC form of the optimal Bayesian predictor.

The finite signal matrix is a subset of the infinite set, which is composed of all possible realizations of the signal wavefronts. If the size of the signal matrix, M , is large enough, the subset could be a good representation of the whole set. Then the postprocessing of the signal matrix could be considered a good approximation to the postprocessing of the original infinite set. The criterion for determining M is problem dependent. Usually, M should be greater than the number of propagating modes, K , so that independent information carried by the propagation modes can be captured by the samples of the signal wavefronts. If the number of elements in the array, N , is smaller than the number of propagating modes, the amount of independent information is limited by N , and so M should be greater than $\min(N, K)$. In this paper, the rule of thumb used was $M \geq 2 \min(N, K)$. This rule has been checked using many simulation examples to show that further increase in M does not significantly change the ROC performance predictions discussed in this paper.

Using the definition of the signal matrix and multiplying both sides of Eq. (1) by the scalar $1/\sqrt{2}\sigma_n$, the hypotheses test simplifies to

$$\begin{aligned} \mathbf{r} &= a\sqrt{\text{SNR}_i}\mathbf{s}_i + \mathbf{n}, \quad i \in 1, \dots, M, \quad \mathbf{n} \sim \mathcal{N}(0, \mathbf{I}_N), \\ H_1: a &\sim \mathcal{N}(0, 1), \quad H_0: a = 0, \end{aligned} \quad (19)$$

where SNR_i is generated by substituting the i th realization of the environmental parameters into Eq. (4). Note that a and \mathbf{n} are normalized so that $\text{Cov}(a) = 1$ and $\text{Cov}(\mathbf{n})$ is an identity matrix of size N . In this discrete hypothesis test, environmental uncertainty is incorporated partly through the probability, $1/M$, of the i th realization of the signal wavefront in the received data and partly through the structure of the signal matrix.

B. Optimal Bayesian predictor

It is often difficult to express the ROC in closed-form for optimal Bayesian detectors. In this section, a useful approximate analytical expression to the ROC is developed.

The development of the optimal Bayesian predictor begins with the discrete likelihood ratio, obtained by utilizing the signal matrix:

$$\begin{aligned} \lambda(\mathbf{r}) &= \sum_{i=1}^M \lambda_i / M, \quad \lambda_i = \exp(B_i |\mathbf{r}^\dagger \mathbf{s}_i|^2) / A_i, \\ A_i &= \text{SNR}_i + 1, \quad B_i = \text{SNR}_i / 2\sigma_n^2 (\text{SNR}_i + 1), \end{aligned} \quad (20)$$

The first approximation is the decision rule approximation. The optimal decision rule “if $\lambda \geq \beta$, decide H_1 and if

$\lambda < \beta$, decide H_0 ” is approximated by “if $\lambda_i \geq \beta$ for any $i \in 1, \dots, M$, decide H_1 and if $\lambda_i < \beta$ for all $i = 1, \dots, M$, decide H_0 .” This approximation has been studied in general.²² However, the authors are not aware of it being used to derive detection performance in a physics-based uncertain environment. Using this approximation, the ROC can be expressed as

$$\begin{aligned} 1 - P_F(\beta) &= P(\lambda_i < \beta, i = 1, \dots, M | H_0), \\ 1 - P_D(\beta) &= P(\lambda_i < \beta, i = 1, \dots, M | H_1). \end{aligned} \quad (21)$$

A second approximation is a constant SNR approximation, i.e., $\text{SNR}_i = \text{SNR}$ for $i = 1, \dots, M$. This assumes that the received signal energy at the receiving array is relatively independent of the perturbations in the ocean transfer function due to ocean environmental and source position uncertainties. For a short vertical array or horizontal array, the constant SNR approximation and the resultant optimal Bayesian predictor might fail because the received signal energy is sensitive to source range variation.

Using this approximation and replacing A_i and B_i with constant A and B , Eq. (21) becomes

$$\begin{aligned} 1 - P_F(\beta') &= P(|\mathbf{r}^\dagger \mathbf{s}_i|^2 < \beta', i = 1, \dots, M | H_0), \\ 1 - P_D(\beta') &= P(|\mathbf{r}^\dagger \mathbf{s}_i|^2 < \beta', i = 1, \dots, M | H_1), \end{aligned} \quad (22)$$

where $\beta' = \log(\beta/A)/B$.

A third approximation is to estimate the rank of the signal matrix as the number of significant eigenvalues of the signal matrix. Assume the rank is R . It is shown²³ that the dimensionality of the data vector in the hypothesis testing problem can be reduced from N to R without affecting the detection performance. Here the hypothesis testing problem differs from Ref. 23 in that the signal vector \mathbf{s}_i is selected from the signal matrix of rank R with the probability $1/M$ whereas it is assumed in the reference that the signal vector is linearly transformed from a parameter vector with a known transform matrix of rank R . However, by considering the signal matrix as the transform matrix and using a column vector as the parameter vector whose i th element is one and whose other elements are zeros, the signal vector \mathbf{s}_i can be treated as a linear transformation of the parameter vector. The procedures in Ref. 23 can be applied here, yielding

$$\begin{aligned} \mathbf{u} &= \sqrt{\text{SNR}}a\mathbf{v}_i + \mathbf{w}, \quad i \in 1, \dots, M, \quad \mathbf{w} \sim \mathcal{N}(0, \mathbf{I}_R), \\ H_1: a &\sim \mathcal{N}(0, 1), \quad H_0: a = 0, \end{aligned} \quad (23)$$

where $\mathbf{u} = \mathbf{U}_1^\dagger \mathbf{r}$, $\mathbf{v}_i = \mathbf{U}_1^\dagger \mathbf{s}_i$, $\mathbf{w} = \mathbf{U}_1^\dagger \mathbf{n}$, and \mathbf{U}_1 is the significant R column of the $N \times N$ unitary matrix \mathbf{U} that is generated from a singular value decomposition of the signal matrix. Note that \mathbf{U}_1 is associated with R significant eigenvalues. The dimensionality of the data vector, the signal wavefront, and the diffuse noise are reduced from N to R .

Since assuming \mathbf{v}_1 as the reference signal wavefront does not change the derivations for P_F and P_D , substituting Eq. (23) in Eq. (22) results in

$$\begin{aligned} 1 - P_F(\beta') &= P(|\mathbf{w}^\dagger \mathbf{v}_1|^2 < \beta', i = 1, \dots, M), \\ 1 - P_D(\beta') &= P(|(\sqrt{\text{SNR}}a\mathbf{v}_1 + \mathbf{w})^\dagger \mathbf{v}_1|^2 < \beta', i = 1, \dots, M). \end{aligned} \quad (24)$$

Both P_D and P_F can be expressed as a multidimensional integral of the probability density function of an independent complex Gaussian vector of size R in a constrained space.

Let $\mathbf{r}_1 = \sqrt{\text{SNR}}a\mathbf{v}_1 + \mathbf{w}$, where \mathbf{r}_1 is a complex Gaussian random vector with covariance matrix $\text{Cov}(\mathbf{r}_1) = 1 + \text{SNR}\mathbf{v}_1\mathbf{v}_1^\dagger$. Using singular value decomposition generates $\text{Cov}(\mathbf{r}_1) = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\dagger$, where \mathbf{Q} is a unitary matrix and $\text{diag}(\mathbf{\Lambda}) = [\text{SNR}+1, 1, \dots, 1]$. Then \mathbf{r}_1 can be transformed to an independent Gaussian vector \mathbf{z} with covariance matrix $\mathbf{\Lambda}$. Substituting $\mathbf{r}_1 = \mathbf{Q}\mathbf{z}$ in Eq. (24) and using $A = \text{SNR} + 1$ yields

$$1 - P_D(\beta') = P(|\mathbf{z}^\dagger \mathbf{Q}^\dagger \mathbf{v}_1|^2 < \beta', i=1, \dots, M)$$

$$= \int_D \exp\left(-|\mathbf{z}_1|^2 A^{-1} - \sum_{i=2}^R |\mathbf{z}_i|^2\right) \times \pi^{-R} A^{-1} d\mathbf{z}, \quad (25)$$

where the domain of the integration is $D = \{|\mathbf{z}^\dagger \mathbf{Q}^\dagger \mathbf{v}_1|^2 < \beta', i=1, \dots, M\}$.

Substituting $\mathbf{w}^\dagger = \mathbf{z}^\dagger \mathbf{Q}^\dagger$ in the P_F expression in Eq. (24) produces

$$1 - P_F(\beta') = \int_D \exp\left(-\sum_{i=1}^R |\mathbf{z}_i|^2\right) \pi^{-R} d\mathbf{z}. \quad (26)$$

Now the domain of the integration for both P_F and P_D is D , which is an irregular convex body in R space. Although accurate P_F and P_D can be obtained with intensive numerical computations, it is difficult to derive closed-form expressions for P_F and P_D because of the irregularity of the integration domain. To obtain an approximate ROC expression, a regular convex body $C = \{|\mathbf{z}_i|^2 < \gamma, i=1, \dots, R\}$ is used to replace D in Eqs. (25) and (26). Considering the integration functions are Gaussian for both P_F and P_D , whose energy is concentrated on the origin, the approximation error due to moderate integration domain differences that are close to the boundary can be moderate. However, it is difficult to derive an analytical error bound. The accuracy of this approximation, in terms of the final goal of obtaining the ROC, is verified in the simulation section.

Regardless of the approximation in the integration domain, the P_F and P_D approximation errors are zero at $\text{SNR}=0$ and $\text{SNR}=\infty$. When the SNR is zero, the P_D expression is the same as the P_F expression, and the parameter γ in domain C can be adjusted such that the integrations on C for P_F and P_D are accurate. When the SNR reaches infinity, the P_D computed using C or D equals to one. The parameter γ in C can be adjusted to compute P_F accurately.

With this approximation, Eqs. (25) and (26) become

$$1 - P_F(\gamma) \approx (1 - \exp(-\gamma))^R,$$

$$1 - P_D(\gamma) \approx (1 - \exp(-\gamma A^{-1}))(1 - \exp(-\gamma))^{R-1}. \quad (27)$$

Finally, combining P_F and P_D yields

$$P_D \approx 1 - (1 - P_F)^{(R-1)/R} \times (1 - (1 - (1 - P_F)^{1/R})^{1/(\text{SNR}+1)}), \quad (28)$$

where the parameter SNR is the signal-to-noise ratio at the receivers [Eq. (4)]. In real applications, the SNR can be re-

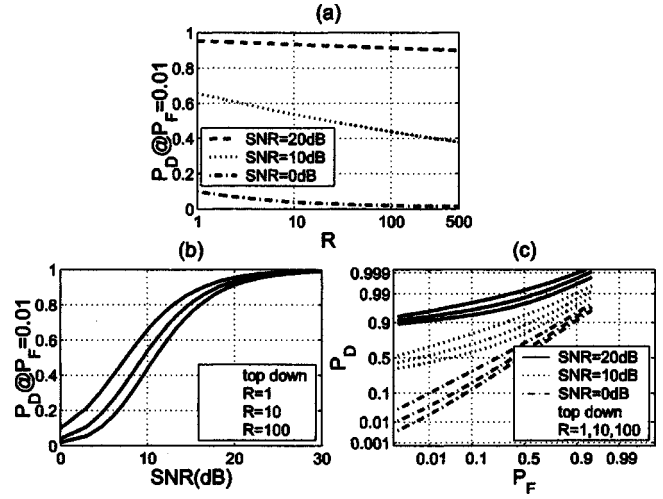


FIG. 2. Analytical detection performance predictions for the optimal Bayesian predictor. (a) P_D as a function of R at fixed $P_F=0.01$, (b) P_D as a function of SNR at fixed $P_F=0.01$, and (c) normal-normal ROC curves as a function of R and SNR.

placed by the average SNR over realizations of the uncertain environment, i.e.,

$$\text{SNR} = \frac{1}{M} \sum_{i=1}^M \text{SNR}_i. \quad (29)$$

The parameter R is the rank of the signal matrix. Strictly speaking, the rank is equal to the number of nonzero eigenvalues. For practical purposes, the rank is determined by the number of significant eigenvalues that exceed a threshold. As shown in Ref. 23, the eigenvalues can be considered as weights to the signal source a . Nonsignificant eigenvalues yield little information (whether $a=0$ or $a \neq 0$) for source detection. Therefore, we only consider the number of significant eigenvalues to estimate the rank in order to predict detection performance. Equation (28) is one of the key results of this paper. The accuracy of this approximate analytical ROC form of the optimal Bayesian predictor is verified in the simulation section.

The computational speed improvement of the optimal Bayesian predictor over the Monte Carlo performance evaluation approach is significant. The primary computational cost for the Monte Carlo approach is the generation of samples of the likelihood ratio for each hypothesis, which is about $O(LMN)$, where L is the number of samples, M is the number of columns in the signal matrix, and N is the number of array sensors. The primary computational cost for the optimal Bayesian predictor is the computation of the eigenvalues, in order to estimate the approximate rank of the signal matrix, which is $O(N^3)$.^{24,25} Typically, M is greater than N , and L is above 1000 to achieve a precision in P_D and P_F of about 0.001, which in turn is about ten times greater than N . Therefore, the analytical approach speedup is typically more than an order of magnitude over the Monte Carlo approach.

Figure 2 illustrates Eq. (28) by plotting P_D as a function of the uncertainty scale R , the SNR at the receivers, and P_F . Figure 2(a) plots P_D as a function of R on a logarithmic scale, assuming $P_F=0.01$ and for SNR=0, 10, or 20 dB. It is shown in Fig. 2(a) that P_D is approximately inversely pro-

portional to R on a logarithmic scale for all three SNR values. Figure 2(b) predicts the optimal detection performance as a function of SNR in the presence of environmental uncertainty of scale R for $P_F=0.01$. The degradation in P_D caused by environmental uncertainty is greater when SNR is less than about 20 dB. In Fig. 2(c), three groups of ROC curves for three SNR values are plotted using normal-normal coordinates. It is often convenient to plot the ROC on normal-normal coordinates. If the underlying distributions of the test statistics for the ROC are normal or Gaussian, the ROC plots as a straight line. In addition, this presentation of the ROC magnifies the region of low probability of false alarm, P_F , similar to log-log coordinates. This plot provides a good comparison of the impact of environmental uncertainty and SNR on the detection performance. It shows that varying R from 1 to 100 or decreasing SNR about 5 dB causes comparable detection performance degradation.

C. Matched-ocean predictor

Since the test statistic for the matched-ocean detector can be considered a chi-square random variable for both hypotheses, an accurate analytical ROC expression for the matched-ocean predictor is available. The resulting form is similar to the ROC for the matched filter detector²³ and the power-type ROC.²⁶

Using \mathbf{s}_k as an abbreviated notation for $\mathbf{s}(\Psi_{\text{known}}, \mathbf{S}_{\text{known}})$ in Eq. (15) results in

$$\begin{aligned} (\lambda'_{\text{matched}}|H_0, \mathbf{s}_k) &= |\mathbf{n}^\dagger \mathbf{s}_k|^2, \\ (\lambda'_{\text{matched}}|H_1, \mathbf{s}_k) &= |(a\mathbf{s}_k + \mathbf{n})^\dagger \mathbf{s}_k|^2. \end{aligned} \quad (30)$$

Let $x_0 = \mathbf{n}^\dagger \mathbf{s}_k$ and $x_1 = (a\mathbf{s}_k + \mathbf{n})^\dagger \mathbf{s}_k$. Since \mathbf{n} is a complex Gaussian vector, scalar a is a complex Gaussian random variable, and \mathbf{s}_k is a known vector, both x_0 and x_1 are complex Gaussian random variables. Their absolute squares are chi-square random variables, with two degrees of freedom. The resultant conditional probability density functions are

$$\begin{aligned} p(\lambda'_{\text{matched}}|H_0, \mathbf{s}_k) &= \exp(\lambda'/C_0)/C_0, \\ p(\lambda'_{\text{matched}}|H_1, \mathbf{s}_k) &= \exp(\lambda'/C_1)/C_1, \end{aligned} \quad (31)$$

where $C_0 = \text{Cov}(x_0) = 2\sigma_n^2 \mathbf{s}_k^\dagger \mathbf{s}_k$ and $C_1 = \text{Cov}(x_1) = C_0(\text{SNR}(\mathbf{s}_k) + 1)$.

Substituting the conditional probability density functions of the test statistics in Eq. (6) produces

$$\begin{aligned} P_F(\beta|\mathbf{s}_k) &= \int_{\beta}^{\infty} d\lambda'_{\text{matched}} p(\lambda'_{\text{matched}}|H_0, \mathbf{s}_k) \\ &= \exp(-\beta/C_0), \\ P_D(\beta|\mathbf{s}_k) &= \int_{\beta}^{\infty} d\lambda'_{\text{matched}} p(\lambda'_{\text{matched}}|H_1, \mathbf{s}_k) \\ &= \exp(-\beta/C_1), \end{aligned} \quad (32)$$

and the ROC expression

$$P_D = P_F^{C_0/C_1} = P_F^{1/(\text{SNR}(\mathbf{s}_k)+1)}. \quad (33)$$

Since the signal wavefront \mathbf{s}_k might be any \mathbf{s}_i in the signal matrix with probability $1/M$, the matched-ocean pre-

dictor is expressed as an average ROC expression

$$P_{D_{\text{matched}}} = \sum_{i=1}^M P_F^{1/(\text{SNR}(\mathbf{s}_i)+1)}/M. \quad (34)$$

This simple matched-ocean predictor is dependent only on the SNR of the known signals.

D. Mean-ocean predictor

To derive the mean-ocean predictor, we assume $\mathbf{r} = a\mathbf{s}_t + \mathbf{n}$ for the H_1 hypothesis, where \mathbf{s}_t represents the real signal wavefront. Replacing $\mathbf{s}(\Psi_{\text{mean}}, \mathbf{S}_{\text{mean}})$ by \mathbf{s}_m for simplicity in Eq. (16) yields the expressions for the test statistic, given \mathbf{s}_m and \mathbf{s}_t :

$$\begin{aligned} (\lambda'_{\text{mean}}|H_0, \mathbf{s}_m) &= |\mathbf{n}^\dagger \mathbf{s}_m|^2, \\ (\lambda'_{\text{mean}}|H_1, \mathbf{s}_t, \mathbf{s}_m) &= |(a\mathbf{s}_t + \mathbf{n})^\dagger \mathbf{s}_m|^2. \end{aligned} \quad (35)$$

Let $x_0 = \mathbf{n}^\dagger \mathbf{s}_m$ and $x_1 = (a\mathbf{s}_t + \mathbf{n})^\dagger \mathbf{s}_m$. Similar to the previous subsection, the absolute squares of both x_0 and x_1 are chi-square random variables, with two degrees of freedom. Their probability density functions are

$$\begin{aligned} p(\lambda'_{\text{mean}}|H_0, \mathbf{s}_m) &= \exp(\lambda'/D_0)/D_0, \\ p(\lambda'_{\text{mean}}|H_1, \mathbf{s}_m, \mathbf{s}_t) &= \exp(\lambda'/D_1)/D_1, \end{aligned} \quad (36)$$

where $D_0 = \text{Cov}(x_0) = 2\sigma_n^2 \mathbf{s}_m^\dagger \mathbf{s}_m$ and $D_1 = \text{Cov}(x_1) = 2\sigma_a^2 |\mathbf{s}_t^\dagger \mathbf{s}_m|^2 + 2\sigma_n^2 |\mathbf{s}_m|^2$.

Given the real signal wavefront \mathbf{s}_t and the mean-ocean signal wavefront \mathbf{s}_m , the analytical ROC form for this mismatched condition can be expressed as

$$P_{D_{\text{mismatched}}} = P_F^{D_0/D_1} = P_F^{1/(\rho(\mathbf{s}_t, \mathbf{s}_m)\text{SNR}(\mathbf{s}_t)+1)}, \quad (37)$$

where $\rho(\mathbf{s}_t, \mathbf{s}_m) = |\mathbf{s}_t^\dagger \mathbf{s}_m|^2 / |\mathbf{s}_t|^2 |\mathbf{s}_m|^2$. The above ROC expression depends on $\text{SNR}(\mathbf{s}_t)$ and $\rho(\mathbf{s}_t, \mathbf{s}_m)$ only, where $\rho(\mathbf{s}_t, \mathbf{s}_m)$ is the correlation of the true signal wavefront with that of the assumed. The parameter ρ characterizes the performance degradation caused by the mismatched environmental parameters in the model.

Considering \mathbf{s}_i might be any \mathbf{s}_i in the signal matrix with probability $1/M$, the ROC for the mean-ocean predictor can be expressed as an average of the mismatched cases:

$$P_{D_{\text{mean}}} = \sum_{i=1}^M P_F^{1/(\rho(\mathbf{s}_i, \mathbf{s}_m)\text{SNR}(\mathbf{s}_i)+1)}/M \quad (38)$$

The mean-ocean predictor predicts the average performance degradation due to environmental mismatch, which is determined primarily by the correlation coefficients $\rho(\mathbf{s}_i, \mathbf{s}_m)$, $i = 1, \dots, M$. It predicts that the performance of the mean-ocean detector decreases as source range increases, as the effects of small errors in the $\exp(-jk_m r_s)$ term [Eq. (2)] accumulate, resulting in mismatch between the mean-ocean signal and the data.

E. Energy predictor

The energy detection algorithm neither requires nor could it use any knowledge about the ocean variables and

source location, even if such information were available. The performance evaluation of the energy detector provides a simple but useful lower performance bound.

To derive the energy predictor, the distributions of the energy test statistic for both hypotheses are required. One special concern in deriving the distribution function for the H_1 hypothesis is that the data depend on the reference signal wavefront \mathbf{s}_i , which might be any column of the signal matrix with probability $1/M$. Another special concern is that the conditional test statistic given the reference signal wavefront, $\lambda'(\mathbf{r}|\mathbf{s}_i)_{\text{energy}} = (\mathbf{a}\mathbf{s}_i + \mathbf{n})^\dagger(\mathbf{a}\mathbf{s}_i + \mathbf{n})$, does not follow a chi-square distribution, as commonly assumed for the energy detector.⁵ The common assumption is made by ignoring data spatial correlations incurred by the presence of the signal.

Since the test statistic is now the sum of correlated gamma random variables, one can obtain analytical forms for the ROC, using the characteristic functions.^{27,28}

The characteristic functions of λ'_{energy} given H_1 and H_0 conditions are

$$\begin{aligned}\Phi(t|H_0) &= E\{\exp(jt\mathbf{r}^\dagger\mathbf{r})|H_0\} = 1/(1 - jt2\sigma_n^2)^N, \\ \Phi(t|H_1) &= \sum_{i=1}^M \Phi(t|H_1, \mathbf{s}_i)/M, \quad (39)\end{aligned}$$

$$\Phi(t|H_1, \mathbf{s}_i) = E\{\exp(jt\mathbf{r}^\dagger\mathbf{r})|H_1, \mathbf{s}_i\} = 1/|I - jtD|,$$

where $D = E\{(\mathbf{a}\mathbf{s}_i + \mathbf{n})(\mathbf{a}\mathbf{s}_i + \mathbf{n})^\dagger\} = 2\sigma_a^2\mathbf{s}_i\mathbf{s}_i^\dagger + 2\sigma_n^2\mathbf{I}$.

The characteristic function $\Phi(t|H_1, \mathbf{s}_i)$ is further decomposed into partial fractions²⁹ to produce Fourier transforms:³⁰

$$\begin{aligned}\Phi(t|H_1, \mathbf{s}_i) &= 1/(1 - jtd_0)^{(N-1)}(1 - jtd_{1i}) \\ &= \sum_{k=1}^{N-1} a_{ik}/(1 - jtd_0)^k + b_i/(1 - jtd_{1i}), \\ d_0 &= 2\sigma_n^2, \quad d_{1i} = d_0A_i, \\ a_{ik} &= (1 - 1/A_i)^{k-N}/A_i, \\ b_i &= -(1 - 1/A_i)^{-N+1}, \quad A_i = \text{SNR}(\mathbf{s}_i) + 1.\end{aligned} \quad (40)$$

The Fourier transforms of the characteristic functions generate the probability density functions of the test statistics:

$$\begin{aligned}p(\lambda'_{\text{energy}}|H_0) &= \gamma(\lambda'_{\text{energy}}/\sigma_n^2, N), \\ p(\lambda'_{\text{energy}}|H_1) &= \sum_{i=1}^M \left(\sum_{k=1}^{N-1} a_{ik} \gamma(\lambda'_{\text{energy}}/d_0, k) \right. \\ &\quad \left. + b_i \gamma(\lambda'_{\text{energy}}/d_{1i}, 1) \right) / M, \quad (41)\end{aligned}$$

$$\gamma(t, N) = t^N \exp(-t) / \Gamma(N).$$

Finally P_D and P_F can be obtained by substituting Eq. (41) in Eq. (6):

$$P_F(\beta) = 1 - \gamma_I(\beta, N),$$

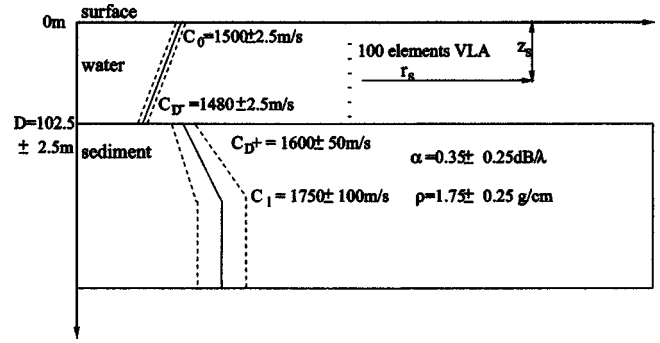


FIG. 3. Benchmark ocean environment.

$$\begin{aligned}P_D(\beta) &= 1 + \sum_{i=1}^M \left(\sum_{k=1}^{N-1} (1 - 1/A_i)^{i-N} \gamma_I(\beta, k) / A_i \right. \\ &\quad \left. - (1 - 1/A_i)^{1-N} \gamma_I(\beta/A_i, 1) \right) / M, \\ A_i &= \text{SNR}_i + 1, \quad (42)\end{aligned}$$

$$\gamma_I(\beta, N) = \int_0^\beta t^{N-1} \exp(-t) dt / \Gamma(N),$$

where γ_I is the incomplete gamma function. The ROC for the energy detector depends on the number of array sensors and the signal-to-noise ratio at the receiving array.

V. RESULTS AND DISCUSSIONS

Simulation scenarios extended from the general mismatch benchmark model (“genlmis”) proposed in the May 1993 NRL Workshop on Acoustic Models in Signal Processing^{17,9,5} were used to check the analytical results. Figure 3 reviews the “genlmis” model, where the vertical line array contains 100 hydrophones spaced 1 m apart ranging from 1 to 100 m in depth. Table II lists the symbols shown in Fig. 3.

In Table II, the ranges of environmental uncertainties ($D, C_0, C_D^-, C_D^+, C_l, \rho, \alpha, z_s$) are denoted using Δ variables. The source range r_s and the upper sediment thickness t were assumed fixed at their mean values in the simulations.

Table III summarizes in four groups a total of 22 different uncertain environmental scenarios and one known environmental scenario. Since source depth D and water depth z_s are two of the most sensitive parameters, their uncertainty

TABLE II. Parameters of uncertain shallow-water environments.

Environmental parameter	Symbol	Value range
Bottom depth	D	102.5 ± ΔD m
Surface sound speed	C ₀	1500 ± ΔC ₀ m/s
Bottom sound speed	C _D ⁻	1480 ± ΔC _D ⁻ m/s
Upper sediment sound speed	C _D ⁺	1600 ± ΔC _D ⁺ m/s
Lower sediment sound speed	C _l	1750 ± ΔC _l m/s
Sediment density	ρ	1.7 ± Δρ g/cm ³
Sediment attenuation	α	0.35 ± Δα dB/A
Sediment thickness	t	100 m
Source depth	z _s	50 ± Δz _s m
Source range	r _s	6 km

TABLE III. Scenarios of environmental uncertainties. Define $\Delta\theta = (\Delta C_0, \Delta C_D^-, \Delta C_D^+, \Delta C_I, \Delta\rho, \Delta\alpha)$ and $\Delta\theta_0 = (2.5, 2.5, 50, 100, 0.25, 0.25)$ for simplicity.

Group of scenarios	Uncertainty configurations
A	$\Delta z_s = \{0, 10, 20, 30, 40, 50\}$
B	$\Delta D = \{0, 1, 2, 3, 4, 5\}$
C	$\Delta z_s = \{0, 10, 20, 30, 40, 50\}, \Delta D = 5, \Delta\theta = \Delta\theta_0$
D	$\Delta D = \{0, 1, 2, 3, 4, 5\}, \Delta\theta = \Delta\theta_0$

ranges, Δz_s and ΔD , are considered separately in groups A and B, respectively. Detection performance appears to be relatively less sensitive to uncertainty in the other environmental parameters taken individually. Therefore, we use $\Delta\theta$ to represent $(\Delta C_0, \Delta C_D^-, \Delta C_D^+, \Delta C_I, \Delta\rho, \Delta\alpha)$ for simplicity and use $\Delta\theta_0 = (2.5, 2.5, 50, 100, 0.25, 0.25)$ as a typical set of values for $\Delta\theta$. In Table III, group A considers six single-source depth-uncertain scenarios with Δz_s increasing from 0 to 50 m. Group B includes six single water-depth-uncertainty scenarios with ΔD increasing from 0 to 5 m. When $\Delta z_s = 0$ in group A or $\Delta D = 0$ in group B, the environmental parameters are configured with their mean values and the environment is known. Group C is group A plus the water depth uncertainty ($\Delta D = 5$ m) and the general ocean uncertainty ($\Delta\theta = \Delta\theta_0$). Group D is group B combined with the presence of the general ocean uncertainty.

A. Performance predictions using the optimal Bayesian predictor

The approximate optimal Bayesian predictor is checked by comparing its prediction results with those obtained using the Monte Carlo performance evaluation approach⁹ for all simulation scenarios listed in Table III.

In the simulations, the input data was generated using Monte Carlo sampling techniques over different ocean environmental, source position, diffuse noise, and signal amplitude and phase realizations based on the data model [Eq. (1)]. In using the Monte Carlo approach, 5000 single snapshot data samples were generated for each hypothesis to produce the samples of the likelihood ratio [Eq. (20)] and to compute P_F and P_D [Eq. (5)]. The ranges of environmental uncertainties and the ratio σ_a^2/σ_n^2 were assumed known *a priori*. Before generating the likelihood ratio samples, the signal matrix \mathfrak{R} was generated using 80 realizations of the signal wavefronts ($M = 80$) and the SNR_i was computed using Eq. (4).

The optimal Bayesian predictor [Eq. (28)] is a function of the approximate rank of the signal matrix and the SNR. Here, the SNR was estimated by computing the mean of the SNR_i 's. The procedure used for choosing an approximate rank is illustrated in Fig. 4(a) for an uncertain scenario. The eigenvalues come from the matrix $\mathfrak{R}^\dagger \mathfrak{R}$. A threshold is used to select the eigenvalues with significant energy, which in turn determines the approximation to the dimensionality of the signal matrix. A “best” threshold is not theoretically tractable. However, the plot of the eigenvalues shows that there is a sharp knee in the curve, which is close to an empirically determined threshold: 5% of the maximum eigenvalue. The

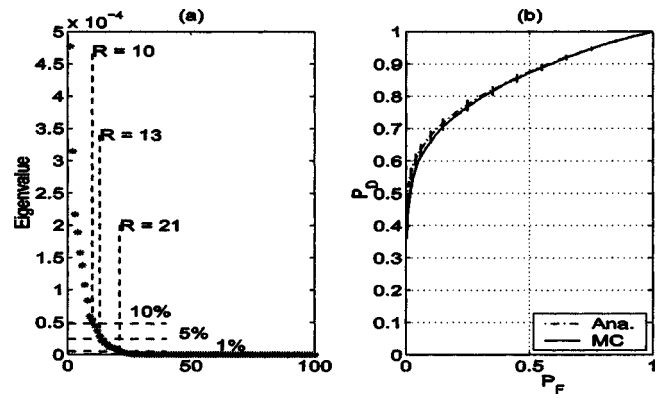


FIG. 4. Analytical detection performance predictions for the optimal Bayesian predictor, for an uncertain environment configured with $\Delta\theta = \Delta\theta_0$ and $\Delta D = 5$. The SNR at the receivers is about 10 dB. (a) Estimation of the approximate rank for the optimal Bayesian predictor, using thresholds of 10%, 5%, and 1% respectively, from left to right. (b) ROC curves for the optimal Bayesian predictor using 5% threshold (dash-dotted curve), 1%–10% thresholds (vertical lines), and the Monte Carlo approach (solid curve).

rank estimated from the empirical threshold and two other thresholds are shown in Fig. 4(a). Figure 4(b) shows that this empirical threshold gives good agreement between Monte Carlo simulation results and analytical ROC results. The variation in P_D due to changes in the threshold from 10% to 1% is illustrated by the short vertical lines in Fig. 4(b). The maximum variation in P_D is less than 0.03. These results illustrate that the ROC is not very sensitive to the threshold value chosen, at and below the knee of the eigenvalue plot.

For all scenarios defined in Table III, it was verified that the ROC results predicted by the optimal Bayesian predictor are consistent with those obtained using the Monte Carlo approach. The ROCs for three of the uncertain scenarios are plotted in Fig. 5 using normal-normal coordinates. These results show that the difference in P_D predicted by the Monte Carlo approach and the optimal Bayesian predictor is small

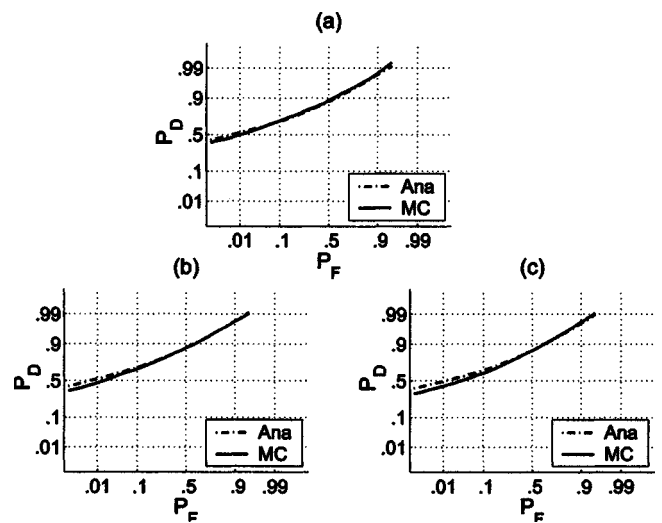


FIG. 5. Comparison of the detection performance predicted by the optimal Bayesian predictor (dash-dotted curve) with the Monte Carlo approach (solid curve) for (a) water depth uncertainty, $\Delta D = 5$, only; (b) water depth uncertainty, $\Delta D = 5$, plus general ocean uncertainty, $\Delta\theta = \Delta\theta_0$; and (c) water depth uncertainty, $\Delta D = 5$, source depth uncertainty, $\Delta z_s = 40$, and general ocean uncertainty, $\Delta\theta = \Delta\theta_0$. The SNR at the receivers is about 10 dB.

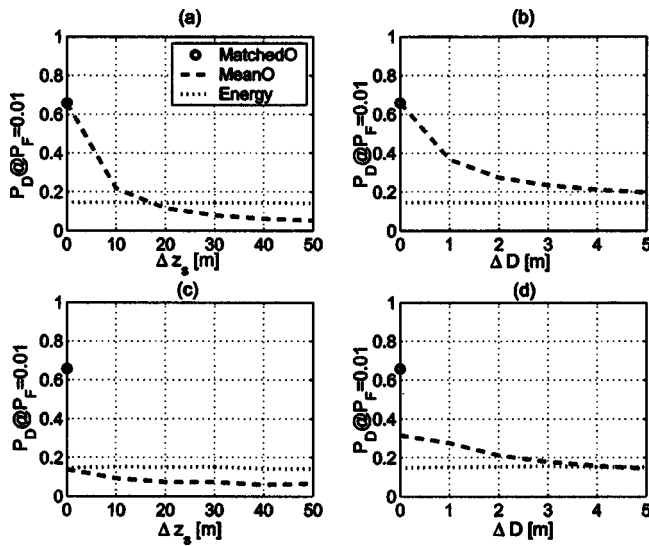


FIG. 6. Comparison of detection performance predictions of matched-ocean, mean-ocean, and energy predictors, plotting probability of correct detection, P_D , as a function of increasing environmental uncertainties. The SNR at the receivers is about 10 dB. Four groups of uncertain scenarios defined in Table III: (a) source depth uncertainty only, (b) water depth uncertainty only, (c) source depth uncertainty plus general ocean uncertainty, and (d) water depth uncertainty plus general ocean uncertainty.

for P_F greater than about 0.1 and is modest (<0.06) for P_F in the range of 0.001–0.1.

B. Quantitative effects of environmental uncertainties

Quantitative effects of environmental uncertainty on detection performance predictions are illustrated in Figs. 6 and 7. In these figures the probability of detection is plotted as a

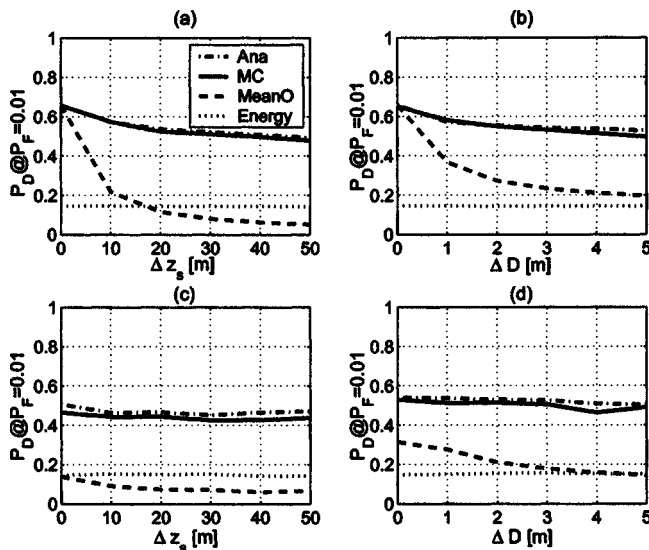


FIG. 7. Comparison of the analytical detection performance prediction, P_D , of the optimal Bayesian predictor (dash-dotted curve), with Monte Carlo evaluation techniques (solid curve) as a function of increasing environmental uncertainties. Also shown for comparison are the performance predictions of the mean-ocean predictor (dashed curve) and the energy predictor (dotted curve). The SNR at the receivers is about 10 dB. Four groups of uncertain scenarios defined in Table III: (a) source depth uncertainty only, (b) water depth uncertainty only, (c) source depth uncertainty plus general ocean uncertainty, and (d) water depth uncertainty plus general ocean uncertainty.

function of the range of environmental uncertainties for the four groups of scenarios defined in Table III.

Figure 6 illustrates the detection performance predicted by the matched-ocean predictor [Eq. (34)], the mean-ocean predictor [Eq. (38)], and the energy predictor [Eq. (42)], which are all precise analytical results. Scenario groups A–D defined in Table III were used to generate Figs. 6(a)–(d), respectively, at a fixed $P_F=0.01$ and for a SNR of about 10 dB at the receivers.

In Figs. 6(a) and (b), the P_D predicted by the matched-ocean predictor, corresponding to $\Delta z_s=0$ and $\Delta D=0$, is determined only by the SNR of the mean-ocean signal and P_F . In Figs. 6(c) and (d), the circles are P_D averaged over the fluctuating SNRs in uncertain environments. The ROCs of the matched-ocean detector in all scenarios provide performance prediction upper bounds for the ideal situation in which there is no environmental uncertainty.

The mean-ocean predictor predicts the same detection performance as that of the matched-ocean predictor when there is no environmental uncertainty. However, the mean-ocean prediction, P_D , falls off rapidly as the range of single parameter uncertainty increases, as shown in Figs. 6(a) and (b). In Figs. 6(c) and (d), the performance predicted by the mean-ocean predictor is even lower with the additional presence of general ocean uncertainty. Although one would expect the performance prediction to fall off as environmental uncertainty increases, it will be shown that it need not fall off this precipitously if the Bayesian optimal predictor is used.

Figures 6(a)–(d) also show P_D predicted by the energy predictor. Its prediction does not change with environmental uncertainties, but remains a small number. This is because the energy detection algorithm does not incorporate the *a priori* knowledge of the environment at all. The performance of the energy predictor provides a simple, yet useful performance lower bound.

Figure 7 illustrates the effects of environmental uncertainty on the optimal Bayesian predictor, at a fixed $P_F=0.01$, for about 10-dB SNR at the receivers. The Bayesian predictor incorporates the degree or range of environmental uncertainty in its prediction. First, this figure shows that the predictions made by the optimal Bayesian predictor agree with those obtained using the Monte Carlo approach for all four groups of scenarios in Table III. Secondly, in contrast to the mean-ocean predictor, the performance of the optimal Bayesian detector falls off more gradually with increasing environmental uncertainty. These results show the importance of fully incorporating environmental uncertainty into the sonar detection performance prediction algorithm.

C. Effects of wrong *a priori* knowledge

Although most of the physics is expressed in the likelihood function, $\lambda(\mathbf{r}|\Psi, \mathbf{S}) = \exp(B(\Psi, \mathbf{S})|\mathbf{r}^\dagger \mathbf{s}(\Psi, \mathbf{S})|^2)/A(\Psi, \mathbf{S})$, the *a priori* knowledge of the uncertain parameters can affect the performance of the Bayesian detector. If the *a priori* distribution of model parameters or the *a priori* ranges of environmental uncertainties are wrong, the Bayesian detection performance degrades, especially if the range of uncertainty of the parameters is underestimated. Figure 8 illustrates the effects of overestimating or underestimating the

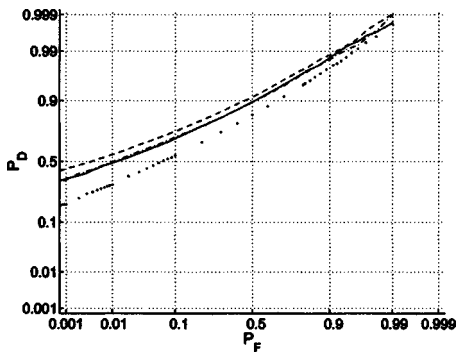


FIG. 8. ROCs illustrating the effect of wrong *a priori* range of ocean depth uncertainty on Bayesian detection performance. From top to bottom, four cases are plotted: assumed and ground truth range of uncertainty are ± 1 m (dashed curve), assumed range of uncertainty is ± 5 m while the truth is ± 1 m (dash-dotted curve), assumed and ground truth range of uncertainty are ± 5 m (solid curve), and assumed range of uncertainty is ± 1 m range while the truth is ± 5 m (dotted curve).

ranges of ocean depth uncertainty on the performance of the Bayesian detector, using the Monte Carlo approach. The number of signal wavefronts is 80 and the simulated data samples for each hypothesis is 20 000.

The ROC curves are plotted using normal-normal coordinates, where the dashed curve denotes the case using a correct assumption of $\Delta D=1$; the dash-dotted curve is for the scenario assuming *a priori* $\Delta D=5$, while the truth is $\Delta D=1$; the solid line denotes the scenario using a correct assumption of $\Delta D=5$; and the dotted one denotes the scenario using an underestimated *a priori* $\Delta D=1$, while the truth is $\Delta D=5$. The maximum degradation from the dashed curve to the dash-dotted curve is 0.03, which shows that if the range of ocean depth uncertainty is overestimated, the performance degradation is modest. However, if the range of uncertainty is underestimated, the P_D performance degradation is as large as 0.1, which is significant. These results indicate that the ROC expression [Eq. (28)] is useful even if the range of uncertainty is overestimated. In summary, these preliminary results indicate that it is preferable to assume that the ocean environmental uncertainty range is somewhat greater than truth.

D. Comparison of performance predictions as a function of source range

In order to get an idea of the impact of ocean environmental uncertainty on detection range, the probability of detection as a function of source range is plotted in Fig. 9 for the general uncertain scenario configured with $\Delta\theta = \Delta\theta_0$ and $\Delta D=5$. From top to bottom, five approaches are illustrated: the matched-ocean predictor, the optimal Bayesian predictor, Monte Carlo performance evaluations for the optimal Bayesian detector, the energy predictor, and the mean-ocean predictor. Although these results are for a fixed source depth, additional simulation results indicate that the general pattern of P_D as a function of source range is relatively independent of source depth.

The difference between the P_D curves predicted by the matched-ocean predictor and the optimal Bayesian predictor reflects the performance degradation due to environmental

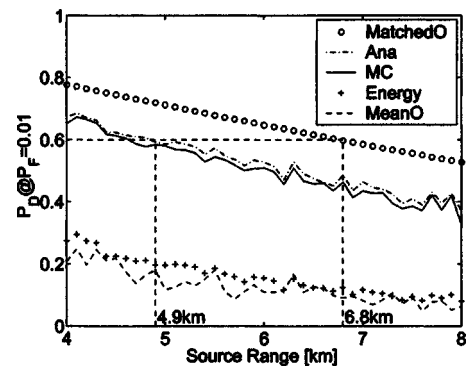


FIG. 9. Detection performance predictions, P_D , as a function of source range in the general ocean uncertainty. Source depth=41 m, $P_F=0.01$, and SNR about 10 dB at source range=6 km. Five performance prediction approaches: the matched-ocean predictor, the optimal Bayesian predictor, Monte Carlo performance evaluation for the optimal Bayesian detector, the energy predictor, and the mean-ocean predictor.

uncertainty, even if one optimally incorporated the environmental uncertainty in the Bayesian prediction algorithm. For example, for $P_D=0.6$, the detection range is in fact 4.9 km, not 6.8 km, as would be predicted by the matched-ocean predictor or classic sonar equation. On the other hand, the mean-ocean and energy predictors give too pessimistic a view of the detection range possible in this particular uncertain ocean environment.

VI. CONCLUSIONS

Analytical approximate ROC expressions, which can be computed rapidly, have been developed for sonar performance prediction. Using a statistical decision theory framework, detection performance prediction algorithms were derived that incorporate the uncertainty of the physics of the acoustic propagation channel as well as the uncertainty of source position in an optimal manner. These analytical expressions for the ROC enable one to compute sonar performance prediction in a much simpler manner than is usually possible using Monte Carlo methods.

Analytical forms were obtained for the ROC for the matched-ocean detector, the mean-ocean detector, and the energy detector. The matched-ocean predictor provides an upper limiting performance bound for the case where there is no uncertainty in the ocean environment. The mean-ocean predictor, which uses only the mean-ocean signal rather than the uncertain environmental knowledge available in the signal matrix, illustrates the degradation in performance due to mismatched model parameters.

The optimal Bayesian predictor incorporates the uncertainty of the physics of the acoustic propagation channel. An algebraic expression was obtained that was shown to be an excellent approximation for the ROC for several examples. This expression indicates that the optimal sonar detection performance (ROC) in diffuse noise depends primarily on the ocean environmental uncertainty, which is captured by the signal matrix, and the mean signal-to-noise ratio at the receivers. These results let one obtain the ROCs in a simple way, as contrasted to computationally intensive Monte Carlo techniques. These results also provide a more meaningful

and realistic performance prediction since it incorporates environmental uncertainty, a feature that is lacking in the classic sonar equation. These results are based on the model that assumes a single source in spatially white noise. However, even in this case, the results can aid in determining how much additional information about the environmental parameters is necessary for a desired improvement in sonar detection performance.

ACKNOWLEDGMENTS

This work was supported by the ONR Uncertainty Defense Research Initiative (DRI). The authors would like to thank Dr. J. L. Krolik for his discussions about the detection model and also thank Dr. S. L. Tantum and Dr. David Schwartz for their discussions about the concept of environmental uncertainty. In addition the authors would like to thank the anonymous reviewers for their many suggestions that helped us improve the presentation of this paper.

¹D. R. Del Balzo, C. Feuillade, and M. M. Rowe, "Effects of water-depth mismatch on matched-field localization in shallow water," *J. Acoust. Soc. Am.* **83**, 2180 (1988).

²A. Tolstoy, "Sensitivity of matched-field processing to sound-speed profile mismatch for vertical arrays in a deep-water pacific environment," *J. Acoust. Soc. Am.* **85**, 2394–2404 (1989).

³H. Schmidt, A. B. Baggeroer, W. A. Kuperman, and E. K. Sheer, "Environmentally tolerant beamforming for high-resolution matched field processing: Deterministic mismatch," *J. Acoust. Soc. Am.* **88**, 1802–1810 (1990).

⁴E. C. Shang, "Environmental mismatching effects on source localization processing in mode space," *J. Acoust. Soc. Am.* **89**, 2285–2290 (1991).

⁵A. Ephraty, J. Tabrikian, and H. Messer, "Underwater source detection using a spatial stationarity test," *J. Acoust. Soc. Am.* **109**, 1053–1063 (2001).

⁶M. D. Collins and W. A. Kuperman, "Focalization: Environmental focusing and source localization," *J. Acoust. Soc. Am.* **90**, 1410–1422 (1991).

⁷J. L. Krolik, "Matched field minimum variance beamforming in a random ocean channel," *J. Acoust. Soc. Am.* **92**, 1408–1419 (1992).

⁸A. M. Richardson and L. W. Nolte, "A *posteriori* probability source localization in an uncertain sound speed, deep ocean environment," *J. Acoust. Soc. Am.* **89**, 2280–2284 (1991).

⁹J. A. Shorey, L. W. Nolte, and J. L. Krolik, "Computationally efficient Monte Carlo estimation algorithm for matched field processing in uncertain ocean environments," *J. Comput. Acoust.* **2**, 285–314 (1994).

¹⁰N. Lee, L. M. Zurk, and J. Ward, "Evaluation of reduced-rank, adaptive matched field processing algorithms for passive sonar detection in a shallow-water environment," *Signals, Systems, and Computers, Record of*

the Thirty-Third Asilomar Conference, Vol. 2, pp. 876–880 (1999).

¹¹J. L. Krolik, "The performance of matched-field beamformers with Mediterranean vertical array data," *IEEE Trans. Signal Process.* **44**, 2605–2611 (1996).

¹²G. B. Smith, H. A. Chandler, and C. Feuillade, "Performance stability of high-resolution matched-field processors to sound-speed mismatch in a shallow-water environment," *J. Acoust. Soc. Am.* **93**, 2617–2626 (1992).

¹³M. Wazenski and D. Alexandrou, "Active, wideband detection and localization in an uncertain multipath environment," *J. Acoust. Soc. Am.* **101**, 1961–1970 (1997).

¹⁴W. J. Seong and S. H. Byun, "Robust matched field-processing algorithm based on feature extraction," *IEEE J. Ocean. Eng.* **27**, 642–652 (2002).

¹⁵S. L. Tantum and L. W. Nolte, "Tracking and localizing a moving source in an uncertain shallow water environment," *J. Acoust. Soc. Am.* **103**, 362–373 (1998).

¹⁶S. L. Tantum, L. W. Nolte, J. L. Krolik, and K. Harmanci, "The performance of matched-field track-before-detect methods using shallow-water pacific data," *J. Acoust. Soc. Am.* **112**, 119–127 (2002).

¹⁷M. B. Porter and A. Tolstoy, "The matched-field processing benchmark problem," *J. Comput. Acoust.* **2**, 161–185 (1994).

¹⁸F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics*, AIP series in Modern Acoustics and Signal Processing (AIP, New York, 1994).

¹⁹M. B. Porter, "The KRAKEN normal mode program" Report No. SM-245, SACLANT Undersea Research Centre, La Spezia, Italy, 1991.

²⁰W. S. Burdick, *Underwater Acoustic System Analysis*, Signal Processing Series (Prentice-Hall, Englewood Cliffs, NJ, 1984).

²¹D. H. Hohnson and D. E. Dudgeon, *Array Signal Processing Concepts and Techniques*, Signal Processing Series (Prentice Hall, Englewood Cliffs, NJ, 1993).

²²L. A. Wainstein and V. D. Zubakov, *Extraction of Signals from Noise* (Prentice-Hall, Englewood Cliffs, NJ, 1962).

²³S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Signal Processing Series (Prentice Hall, Englewood Cliffs, NJ, 1998).

²⁴Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, H. van der Vorst, and editors, *Templates for the Solution of Algebraic Eigenvalue Problems: a Practical Guide* (SIAM, Philadelphia, 2000).

²⁵T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms* (McGraw-Hill, New York, MIT, Cambridge, 1998).

²⁶T. G. Birdsall, "The theory of signal detectability: Roc curves and their character," *Cooley Electron. Lab., Univ. Michigan, Ann Arbor, Tech. Rep. Vol. 177*, 1973.

²⁷P. R. Krishnaiah and M. M. Rao, "Remarks on a multivariate gamma distribution," *Am. Math. Monthly* **68**, 342–346 (1961).

²⁸V. A. Aalo, "Performance of Maximal-Ratio diversity systems in a correlated Nakagami-Fading environment," *IEEE Trans. Commun.* **43**, 2360–2370 (1995).

²⁹I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products, Corrected and Enlarged Edition* (Academic, New York, 1980).

³⁰G. A. Campbell and R. D. Foster, *Fourier Integrals for Practical Applications* (Van Nostrand, New York, 1942).

Bayesian sonar detection performance prediction in the presence of interference in uncertain environments

Liewei Sha and Loren W. Nolte^{a)}

Department of Electrical and Computer Engineering, Box 90291, Duke University,
Durham, North Carolina 27708-0291

(Received 18 April 2004; revised 21 January 2005; accepted 25 January 2005)

The detection performance of sonar systems can be greatly limited by the presence of interference and environmental uncertainty. The classic sonar equation does not take into account these two limiting factors and is inaccurate in predicting sonar detection performance. Here we have developed closed-form receiver operating characteristic (ROC) performance expressions for the Bayesian detector in the presence of interference in uncertain environments. Various scenarios extended from a NRL benchmark shallow-water model were used to test the analytical ROC expressions and to analyze the effects of interference and environmental uncertainty on detection performance. The results show that (1) the degradation on detection performance due to interference is greatly magnified by the presence of environmental uncertainty; (2) Bayesian sonar detection performance depends on the following fundamental parameters: the signal-to-noise ratio, the rank of the signal matrix, and the signal-to-interference coefficient; (3) the proposed analytical ROC performance predictions can be computed much faster than performance evaluations with commonly used Monte Carlo techniques. © 2005 Acoustical Society of America.

[DOI: 10.1121/1.1871732]

PACS numbers: 43.30.Wi, 43.60.Gk, 43.60.Uv [AIT]

Pages: 1954–1964

I. INTRODUCTION

Sonar detection performance prediction has received significant attention by the underwater acoustic community. In many sonar systems,¹ the sonar equation is used as a means for predicting the detection performance. The classic sonar equation assumes that both the signal wavefront and the noise field directionality are known *a priori*. However, under practical circumstances, the signal wavefront and noise field directionality are usually uncertain. The signal wavefront is a function of the ocean environment, the source position, and the receiving array configuration. Ocean environmental parameters such as water depth, sound speed profile, bottom density, etc., have strong spatial and temporal variability, whose values are not known precisely. Source motion limits the accuracy of the estimate of the source position. Uncertainties in both the source position and ocean environmental parameters translate to uncertainty in the signal wavefront. The noise field directionality is usually uncertain due to the presence of interferers with uncertain locations in the uncertain ocean environment. In summary, the classic sonar equation does not take into account environmental uncertainty and hence is inaccurate, often too optimistic for a sonar detection performance prediction in a realistic circumstance.

The sensitivity of the localization algorithms to the uncertain ocean environmental parameters and the uncertain source position parameters has generated interest in the research community.^{2,3} Several localization algorithms that are robust to these uncertainties have been proposed, such as the minimum variance beamformer with multiple neighboring

location constraints (MV_NCL),³ the optimum uncertain field processor (OUFP),^{4,5} the minimum variance beamformer with environmental perturbation constraints (MV_EPC),⁶ and the reduced-rank MFP algorithms.⁷ With a few modifications, these algorithms can be used to detect the presence of a target. Robust algorithms directly proposed for detection problems include the stationarity test⁸ and the Bayesian detectors.^{5,9} The detection performance of those algorithms have been reported.^{5,7,8} For example, the detection performance of reduced-rank MFP algorithms has been tested using an array gain metric and using data from the Santa Barbara Channel experiment.⁷ The spatial stationarity test⁸ has been evaluated using Mediterranean vertical array data.¹⁰ However, those performance evaluations do not provide a fundamental relationship between environmental uncertainty and detection performance. Further, most performance evaluations^{5,9,11} are computed using Monte Carlo techniques rather than through analytical means, which is computationally intensive.

The presence of interference is another limiting factor to detection performance besides environmental uncertainty. Some detection algorithms assume a known plane wave interference model.^{12,13} The resultant performance analysis can be too optimistic. Other detection algorithms^{14,15} assume that the noise field is completely unknown, requiring either *in situ* measurements or adaptive estimations from noise-only data, which do not take advantage of available information about the environment and the interferers.

Here we develop a closed-form ROC expression for the Bayesian detector in the presence of interference in uncertain environments, which is different from previous work¹⁶ that does not take into account interference. The principal cause of the interference modeled here are the acoustic sources

^{a)}Electronic mail: lwn@ee.duke.edu

other than the target. The Bayesian detector is modified from the OUPF,⁴ which incorporates the *a priori* information of the interferers. The resultant ROC expression is an algebraic function of a few parameters: the signal-to-interference coefficient, which captures the effect of interference in uncertain environments; the rank of the signal matrix, which characterizes the degree of environmental uncertainty; and the signal-to-noise ratio, which characterizes the combined effect of the source level, the noise level, the propagation loss, and array gain. Simulated data generated using a NRL benchmark shallow-water model^{8,5,17} are used to verify the proposed Bayesian detection performance prediction approach, by comparing analytical results with Monte Carlo performance evaluation results. Good agreement of those results demonstrates that the simple analytical ROC expression derived here captures fundamental parameters of the performance of the Bayesian detectors. This provides a more realistic detection performance prediction than does the classic sonar equation. The simulation results also indicate that detection performance degradation due to interference is magnified by the presence of environmental uncertainty.

The paper is organized as follows. In Sec. II, the signal and interference model is defined. In Sec. III, the detection problem is formulated. In Sec. IV, the Bayesian detector that incorporates the *a priori* information of the interferers and uncertain environmental parameters is presented. In Secs. V, VI, and VII, analytical ROC expressions are developed for the Bayesian detector in known environments and in uncertain environments. In Sec. VIII we verify the proposed performance prediction approach and the effect of interference on detection performance using simulated data.

II. SIGNAL AND INTERFERENCE MODEL

Typical interference models include a conventional plane wave model for detection and parameter estimation problems,^{12,13} and the Gaussian distribution model in beamforming.¹⁸ Here, the uncertain signal and interference at the receivers are both expressed as a function of the uncertain ocean environmental parameters and acoustic source (signal source or interferer) position parameters. With this translation of uncertainties, the model enables a detection algorithm to incorporate the *a priori* knowledge at the ocean and source parameter level, rather than at the signal or interference wavefront level. The advantage of this translation is that it is much easier to get a stochastic description of the ocean and source parameters than to obtain the *a priori* probability density function for the signal and interference wavefronts.

Ocean and source position parameters are defined along with their stochastic descriptions. The ocean environmental parameters, Ψ , represent ocean properties affecting the acoustic propagation. Such properties include water depth, sound speed profile, surface roughness, bottom composition and roughness, etc. The signal source position relative to the receiving array is denoted by \mathbf{S} . The number of interferers, K , is assumed a known constant. The k th interferer position is denoted by \mathbf{S}_k and the set of all interferer positions is denoted by $\bar{\mathbf{S}}_f$. The set of parameters Ψ , \mathbf{S} , and $\bar{\mathbf{S}}_f$ together

constitute an uncertain environment. Environmental uncertainty is quantified by assigning an *a priori* probability density function (pdf) over a finite range of possible values. It is assumed that Ψ , \mathbf{S} , and $\bar{\mathbf{S}}_f$ are statistically independent *a priori*, with uniform pdf's denoted by $p(\Psi)$, $p(\mathbf{S})$, and $p(\bar{\mathbf{S}}_f)$ that are defined on uncertainty ranges Ω_Ψ , $\Omega_{\mathbf{S}}$, and $\Omega_{\bar{\mathbf{S}}_f}$. The uncertainty ranges are assumed known.

The received signal and interference are both assumed to be a spatial vector in the frequency domain, narrow band, centered at a known frequency f_0 . The received signal consists of three components: the signal wavefront, the signal source, and the received signal energy. The normalized signal wavefront $\mathbf{s}(\Psi, \mathbf{S})$ is given by

$$\mathbf{s}(\Psi, \mathbf{S}) = \mathbf{H}(\Psi, \mathbf{S}) / \|\mathbf{H}(\Psi, \mathbf{S})\|, \quad (1)$$

where $\mathbf{H}(\Psi, \mathbf{S})$ is the frequency-domain ocean acoustic transfer function, which is the acoustic pressure field sampled at an arbitrary receiving array of N sensors, given ocean environmental parameters Ψ and source position parameters \mathbf{S} . Here $\|\cdot\|$ means the 2-norm on a complex vector space.

The signal source a represents uncertainties in the amplitude and phase of the source in the frequency domain. It is a normalized random variable with zero mean and unit variance, $a \sim \mathcal{N}(0, 1)$. Here \sim means "is distributed as" and \mathcal{N} means complex Gaussian distribution.

Because the effects of the signal variance, σ_a^2 , the diffuse noise variance, σ_n^2 , and the norm of the ocean transfer function on the detection performance are indistinguishable, a single model parameter, the signal-to-noise ratio (SNR), is defined as

$$\text{SNR}(\Psi, \mathbf{S}) = \sigma_a^2 \|\mathbf{H}(\Psi, \mathbf{S})\|^2 / \sigma_n^2. \quad (2)$$

It is not an element-level SNR but an array-level SNR. The ratio σ_a^2 / σ_n^2 is assumed a known constant. The $\text{SNR}(\Psi, \mathbf{S})$ depends on σ_a^2 / σ_n^2 and the realizations of the signal wavefront due to uncertain environmental parameters. In practice, an estimate of the noise variance may be obtained using additional measurements at neighboring frequencies that contain only diffuse noise. The received signal energy, $\sigma_a^2 \|\mathbf{H}(\Psi, \mathbf{S})\|^2$, may be estimated from measurements at the source frequencies, and consequently the SNR may be obtained. Using the definition of the SNR, the received signal is expressed as $\sqrt{\text{SNR}(\Psi, \mathbf{S})} \sigma_n^2 a \mathbf{s}(\Psi, \mathbf{S})$.

The k th interference, $\sqrt{\text{INR}_k(\Psi, \mathbf{S}_k)} \sigma_n^2 b_k \mathbf{f}_k(\Psi, \mathbf{S}_k)$, also consists of three components. The normalized interference wavefront is given by

$$\mathbf{f}_k(\Psi, \mathbf{S}_k) = \mathbf{H}(\Psi, \mathbf{S}_k) / \|\mathbf{H}(\Psi, \mathbf{S}_k)\|. \quad (3)$$

The interference source, $b_k \sim \mathcal{N}(0, 1)$, represents normalized interferer amplitude and phase. The energy of the k th interference is given by $\text{INR}_k(\Psi, \mathbf{S}_k) \sigma_n^2$, where the array level interference-to-noise ratio (INR) depends on the source variance of the k th interferer, $\sigma_{b_k}^2$, the diffuse noise variance, and the norm of the ocean transfer function:

$$\text{INR}_k = \sigma_{b_k}^2 \|\mathbf{H}(\Psi, \mathbf{S}_k)\|^2 / \sigma_n^2. \quad (4)$$

Here, the ratio $\sigma_{b_k}^2/\sigma_n^2$ is assumed a known constant. In practice, an estimate of σ_{b_k} may be obtained from measurements at neighboring interference frequencies other than the signal frequency.

III. DETECTION PROBLEM

The detection problem is formulated as a doubly composite binary hypotheses testing problem. For an arbitrary array of N sensors, the observation is an $N \times 1$ spatial vector in the frequency domain, obtained using a narrow-band Fourier transform of the snapshot. The i th data element is transformed from the snapshot received by the sensor at the i th location. Given the observation, we must decide between two hypotheses: the “null” hypothesis H_0 and the “signal present” hypothesis H_1 , in the presence of interference and environmental uncertainty. The null hypothesis assumes that the data consist of diffuse noise and uncertain interference. The H_1 hypothesis assumes that the data consists of the received signal, in addition to diffuse noise and interference.

$$\begin{aligned} H_1: \mathbf{r} &= \sqrt{\text{SNR}(\boldsymbol{\Psi}, \mathbf{S})} \sigma_n^2 a \mathbf{s}(\boldsymbol{\Psi}, \mathbf{S}) + \mathbf{n}_1, \\ H_0: \mathbf{r} &= \mathbf{n}_1, \\ \mathbf{n}_1 &= \sum_{k=1}^K \sqrt{\text{INR}_k(\boldsymbol{\Psi}, \mathbf{S}_k)} \sigma_n^2 b_k \mathbf{f}_k(\boldsymbol{\Psi}, \mathbf{S}_k) + \sigma_n \mathbf{n}_0, \\ \mathbf{n}_0 &\sim \mathcal{N}(0, \mathbf{I}_N), \\ \boldsymbol{\Psi} &\sim p(\boldsymbol{\Psi}), \quad \mathbf{S} \sim p(\mathbf{S}), \quad \mathbf{S}_k \sim p(\mathbf{S}_k), \end{aligned} \quad (5)$$

where diffuse noise $\sigma_n \mathbf{n}_0$ is modeled by a spatial complex Gaussian vector in the frequency domain.

In Eq. (5), the signal wavefront is a function of the continuous random variables $\boldsymbol{\Psi}$ and \mathbf{S} . To computationally predict the performance of detection algorithms in an uncertain environment, a discrete representation of the uncertain signal wavefront is needed. A matrix composed of M realizations of the signal wavefront due to environmental uncertainties, referred to as the signal matrix \mathfrak{R} , is defined by applying Monte Carlo sampling techniques,¹⁶

$$\begin{aligned} \mathfrak{R} &= [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_M], \\ &= [\mathbf{s}(\boldsymbol{\Psi}, \mathbf{S})_1, \mathbf{s}(\boldsymbol{\Psi}, \mathbf{S})_2, \dots, \mathbf{s}(\boldsymbol{\Psi}, \mathbf{S})_M], \end{aligned} \quad (6)$$

where $(\boldsymbol{\Psi}, \mathbf{S})_i (i = 1 \dots M)$ represents the i th realization of the environmental parameters $(\boldsymbol{\Psi}, \mathbf{S})$. If M is large enough, the signal matrix is a good representation of realizations of uncertain signal wavefronts. Therefore, the detection problem can be formulated based on the signal matrix rather than on a continuous form of the signal wavefronts.

The k th interference matrix is defined similarly to the signal matrix:

$$\begin{aligned} T_k &= [\mathbf{f}_{k1}, \mathbf{f}_{k2}, \dots, \mathbf{f}_{kL}], \\ &= [\mathbf{f}_k((\boldsymbol{\Psi}, \mathbf{S}_k)_1), \mathbf{f}_k((\boldsymbol{\Psi}, \mathbf{S}_k)_2), \dots, \mathbf{f}_k((\boldsymbol{\Psi}, \mathbf{S}_k)_L)]. \end{aligned} \quad (7)$$

Although the signal and interference both involve the same uncertain environment $(\boldsymbol{\Psi})$, their source position uncertainties $(\mathbf{S}$ vs $\mathbf{S}_k)$ may differ. The number of realizations for the interference is denoted by L .

By applying the concept of the signal and interference matrices, and multiplying by the scalar $1/\sigma_n$ on both sides of Eq. (5) for simplicity, the detection problem is formulated in a discrete form:

$$\begin{aligned} H_1: \mathbf{r} &= \sqrt{\text{SNR}_i} a \mathbf{s}_i + \mathbf{n}_1, \quad i \in 1 \dots M \\ H_0: \mathbf{r} &= \mathbf{n}_1, \\ \mathbf{n}_1 &= \sum_{k=1}^K \sqrt{\text{INR}_{kl}} b_k \mathbf{f}_{kl} + \mathbf{n}_0, \quad l \in \dots L \\ \mathbf{n}_0 &\sim \mathcal{N}(0, \mathbf{I}_N), \end{aligned} \quad (8)$$

where $\text{SNR}_i = \text{SNR}((\boldsymbol{\Psi}, \mathbf{S})_i)$, $\text{INR}_{kl} = \text{INR}_k((\boldsymbol{\Psi}, \mathbf{S}_k)_l)$, \mathbf{s}_i is an arbitrary column of the signal matrix with probability $1/M$, and \mathbf{f}_{kl} is an arbitrary column of the k th interference matrix with probability $1/L$.

In the detection problem, diffuse noise (\mathbf{n}_0) , interference source (b_k) , signal source (a) , and environmental parameters $(\boldsymbol{\Psi}, \mathbf{S}$, and $\mathbf{S}_k)$ are assumed statistically independent.

IV. BAYESIAN DETECTOR

From signal detection theory, the optimal detector is the Bayesian detector that fully incorporates the *a priori* knowledge of the uncertain parameters. The derivation of the Bayesian detector begins with the likelihood ratio²⁰

$$\lambda(\mathbf{r}) = \frac{p(\mathbf{r}|H_1)}{p(\mathbf{r}|H_0)}. \quad (9)$$

Based on the signal and interference model, the data vector under both hypotheses are assumed to be conditional complex Gaussian:

$$\begin{aligned} p(\mathbf{r}|H_1, \boldsymbol{\Psi}, \bar{\mathbf{S}}_f, \mathbf{S}) \\ = \exp(-\mathbf{r}^\dagger \mathbf{M}_1^{-1}(\boldsymbol{\Psi}, \bar{\mathbf{S}}_f, \mathbf{S}) \mathbf{r}) / |\pi \mathbf{M}_1(\boldsymbol{\Psi}, \bar{\mathbf{S}}_f, \mathbf{S})|, \end{aligned} \quad (10)$$

$$p(\mathbf{r}|H_0, \boldsymbol{\Psi}, \bar{\mathbf{S}}_f) = \exp(-\mathbf{r}^\dagger \mathbf{M}_0^{-1}(\boldsymbol{\Psi}, \bar{\mathbf{S}}_f) \mathbf{r}) / |\pi \mathbf{M}_0(\boldsymbol{\Psi}, \bar{\mathbf{S}}_f)|,$$

where

$$\mathbf{M}_1(\boldsymbol{\Psi}, \bar{\mathbf{S}}_f, \mathbf{S}) = \text{SNR}(\boldsymbol{\Psi}, \mathbf{S}) \mathbf{s}(\boldsymbol{\Psi}, \mathbf{S}) \mathbf{s}^\dagger(\boldsymbol{\Psi}, \mathbf{S}) + \mathbf{M}_0(\boldsymbol{\Psi}, \bar{\mathbf{S}}_f), \quad (11)$$

$$\mathbf{M}_0(\boldsymbol{\Psi}, \bar{\mathbf{S}}_f) = \sum_{k=1}^K \text{INR}_k(\boldsymbol{\Psi}, \mathbf{S}_k) \mathbf{f}_k(\boldsymbol{\Psi}, \mathbf{S}_k) \mathbf{f}_k^\dagger(\boldsymbol{\Psi}, \mathbf{S}_k) + \mathbf{I}_N.$$

The likelihood ratio [Eq. (9)] is a doubly composite hypotheses test:

$$\begin{aligned} \lambda(\mathbf{r}) \\ = \frac{\int \boldsymbol{\Psi} \int \bar{\mathbf{S}}_f \int \mathbf{S} d\boldsymbol{\Psi} d\bar{\mathbf{S}}_f d\mathbf{S} p(\mathbf{r}|H_1, \boldsymbol{\Psi}, \bar{\mathbf{S}}_f, \mathbf{S}) p(\boldsymbol{\Psi}) p(\bar{\mathbf{S}}_f) p(\mathbf{S})}{\int \boldsymbol{\Psi} \int \bar{\mathbf{S}}_f d\boldsymbol{\Psi} d\bar{\mathbf{S}}_f p(\mathbf{r}|H_0, \boldsymbol{\Psi}, \bar{\mathbf{S}}_f) p(\boldsymbol{\Psi}) p(\bar{\mathbf{S}}_f)}. \end{aligned} \quad (12)$$

The probability density functions of the data vector with environmental and source position uncertainties for H_1 and H_0 hypotheses are given by the nominator and the denominator of the right side of Eq. (12), respectively. Both are non-Gaussian and are the result of integrating the conditional complex Gaussian probability density functions over the uncertain environmental parameters.

In order to capture analytically the detection performance in a tractable manner a simpler approximate test statistic is postulated. The approximation is that the noise field is a multivariate complex Gaussian distribution, characterized by the covariance matrix $\hat{\mathbf{M}}_0$, where $\hat{\mathbf{M}}_0$ is $\mathbf{M}_0(\boldsymbol{\Psi}, \bar{\mathbf{S}}_f)$ [Eq. (11)] integrated over uncertain ocean environmental parameters and interferer positions:

$$\begin{aligned}\hat{\mathbf{M}}_0 &= \int_{\boldsymbol{\Psi}} \int_{\bar{\mathbf{S}}_f} d\boldsymbol{\Psi} d\bar{\mathbf{S}}_f p(\boldsymbol{\Psi}) p(\bar{\mathbf{S}}_f) \mathbf{M}_0(\boldsymbol{\Psi}, \bar{\mathbf{S}}_f) \\ &= \int_{\boldsymbol{\Psi}} \int_{\bar{\mathbf{S}}_f} d\boldsymbol{\Psi} d\bar{\mathbf{S}}_f p(\boldsymbol{\Psi}) p(\bar{\mathbf{S}}_f) \\ &\quad \times \sum_{k=1}^K \text{INR}_k(\boldsymbol{\Psi}, \mathbf{S}_k) \mathbf{f}_k(\boldsymbol{\Psi}, \mathbf{S}_k) \mathbf{f}_k^\dagger(\boldsymbol{\Psi}, \mathbf{S}_k) + \mathbf{I}.\end{aligned}\quad (13)$$

The covariance matrix $\hat{\mathbf{M}}_0$ incorporates the effect of environmental uncertainty and the presence of interference through integrations over uncertain ocean environmental parameters and interferer position parameters. The matrix $\hat{\mathbf{M}}_0$ is no longer an explicit function of interferer positions and environmental parameters, but a function of interference and environmental uncertainty ranges. This approach is different from the approach that assumes a known covariance matrix, which requires too much *a priori* information. It also differs from the approach that uses the estimation of the covariance matrix from the collection of noise-only data, which relies too much on the data, but does not take advantage of the available *a priori* information about the interferer positions and the environmental parameters.

The concept of the interference matrices is applied to numerically compute the covariance matrix, yielding

$$\hat{\mathbf{M}}_0 = \frac{1}{L} \sum_{l=1}^L \sum_{k=1}^K \text{INR}_{kl} \mathbf{f}_{kl} \mathbf{f}_{kl}^\dagger + \mathbf{I}.\quad (14)$$

Assuming that the noise field is multivariate complex Gaussian, the distributions of the data conditional to both hypotheses become

$$\begin{aligned}p(\mathbf{r}|H_1, \boldsymbol{\Psi}, \mathbf{S}) &= \exp(-\mathbf{r}^\dagger \hat{\mathbf{M}}_1^{-1} \mathbf{r}) / |\pi \hat{\mathbf{M}}_1|, \\ p(\mathbf{r}|H_0) &= \exp(-\mathbf{r}^\dagger \hat{\mathbf{M}}_0^{-1} \mathbf{r}) / |\pi \hat{\mathbf{M}}_0|,\end{aligned}\quad (15)$$

where $\hat{\mathbf{M}}_1 = \text{SNR}(\boldsymbol{\Psi}, \mathbf{S}) \mathbf{s}(\boldsymbol{\Psi}, \mathbf{S}) \mathbf{s}(\boldsymbol{\Psi}, \mathbf{S})^\dagger + \hat{\mathbf{M}}_0$.

Therefore the likelihood ratio can be approximated by

$$\tilde{\lambda}(\mathbf{r}) = \int_{\boldsymbol{\Psi}} \int_{\mathbf{S}} d\boldsymbol{\Psi} d\mathbf{S} \lambda(\mathbf{r}|\boldsymbol{\Psi}, \mathbf{S}) p(\boldsymbol{\Psi}) p(\mathbf{S}),\quad (16)$$

where the conditional likelihood ratio is given by

$$\lambda(\mathbf{r}|\boldsymbol{\Psi}, \mathbf{S}) = p(\mathbf{r}|H_1, \boldsymbol{\Psi}, \mathbf{S}) / p(\mathbf{r}|H_0).\quad (17)$$

Using Woodbury's identity,²¹

$$\hat{\mathbf{M}}_1^{-1} = \hat{\mathbf{M}}_0^{-1} - \frac{\hat{\mathbf{M}}_0^{-1} \text{SNR}(\boldsymbol{\Psi}, \mathbf{S}) \mathbf{s}(\boldsymbol{\Psi}, \mathbf{S}) \mathbf{s}(\boldsymbol{\Psi}, \mathbf{S})^\dagger \hat{\mathbf{M}}_0^{-1}}{1 + \text{SNR}(\boldsymbol{\Psi}, \mathbf{S}) \mathbf{s}(\boldsymbol{\Psi}, \mathbf{S})^\dagger \hat{\mathbf{M}}_0^{-1} \mathbf{s}(\boldsymbol{\Psi}, \mathbf{S})}.\quad (18)$$

Using properties of the matrix determinant,

$$|\hat{\mathbf{M}}_1| = |\hat{\mathbf{M}}_0| (1 + \text{SNR}(\boldsymbol{\Psi}, \mathbf{S}) \mathbf{s}(\boldsymbol{\Psi}, \mathbf{S})^\dagger \hat{\mathbf{M}}_0^{-1} \mathbf{s}(\boldsymbol{\Psi}, \mathbf{S})).\quad (19)$$

Substituting Eqs. (15), (18), and (19) in Eq. (17) results in

$$\begin{aligned}\lambda(\mathbf{r}|\boldsymbol{\Psi}, \mathbf{S}) &= \frac{1}{1 + \text{SNR}(\boldsymbol{\Psi}, \mathbf{S}) G(\boldsymbol{\Psi}, \mathbf{S})} \\ &\quad \times \exp\left(-\frac{\text{SNR}(\boldsymbol{\Psi}, \mathbf{S}) |\mathbf{r}^\dagger \hat{\mathbf{M}}_0^{-1} \mathbf{s}(\boldsymbol{\Psi}, \mathbf{S})|^2}{1 + \text{SNR}(\boldsymbol{\Psi}, \mathbf{S}) G(\boldsymbol{\Psi}, \mathbf{S})}\right), \\ G(\boldsymbol{\Psi}, \mathbf{S}) &= \mathbf{s}^\dagger(\boldsymbol{\Psi}, \mathbf{S}) \hat{\mathbf{M}}_0^{-1} \mathbf{s}(\boldsymbol{\Psi}, \mathbf{S}),\end{aligned}\quad (20)$$

where the term $G(\boldsymbol{\Psi}, \mathbf{S})$ is referred to as the signal-to-interference coefficient. It includes uncertain interferer position parameters through the approximate covariance matrix of the noise field, $\hat{\mathbf{M}}_0$. The product of SNR and G can be referred to as the "signal-to-interference-plus-noise ratio (SINR)," which is comparable to the SINR defined in the literature.^{21,13} The difference between our definition of SINR and others is that the product of the SNR and G in Eq. (20) is derived from a Bayesian detection viewpoint. Other definitions of SINR are results from a beamforming perspective, concentrating on the total effect of the noise field, usually in a known environment.

Finally, the Bayesian detector that incorporates the *a priori* information of environmental uncertainties and the interferers is given by Eqs. (16), (20), and (13).

To numerically compute the likelihood ratio, the concept of the signal matrix is applied to generate a discrete version of the approximate likelihood ratio:

$$\begin{aligned}\tilde{\lambda}(\mathbf{r}) &= \frac{1}{M} \sum_{i=1}^M \lambda_i, \\ \lambda_i &= \frac{1}{1 + \text{SNR}_i G_i} \exp\left(-\frac{\text{SNR}_i |\mathbf{r}^\dagger \hat{\mathbf{M}}_0^{-1} \mathbf{s}_i|^2}{1 + \text{SNR}_i G_i}\right), \\ G_i &= \mathbf{s}_i^\dagger \hat{\mathbf{M}}_0^{-1} \mathbf{s}_i, \\ \text{SNR}_i &= \text{SNR}((\boldsymbol{\Psi}, \mathbf{S})_i).\end{aligned}\quad (21)$$

Note that the computation of G_i includes the inverse of $\hat{\mathbf{M}}_0$, which is sensitive to small eigenvalues of $\hat{\mathbf{M}}_0$. In the presence of small eigenvalues, an "eigenvalue threshold" method¹⁹ can be used to compute the matrix inversion.

V. PERFORMANCE PREDICTIONS IN KNOWN ENVIRONMENTS

In a known environment, it is assumed that the ocean and the source position parameters are known *a priori*. Consequently, the signal and interference wavefronts are assumed known. In this case, the ROC for the proposed Bayesian detector provides a benchmark to be compared with detection performance predictions in uncertain environments. Here, the analytical ROC performance prediction expressions are derived. The effect of interference on the detection performance in known environments is analyzed.

A known environment assumption is equivalent to assuming that the *a priori* distributions $p(\boldsymbol{\Psi})$, $p(\mathbf{S})$, and $p(\mathbf{S}_k)$ are delta functions. Substituting these conditions in Eq. (20) results in the likelihood ratio specific to a known environment:

$$\lambda(\mathbf{r}) = \frac{1}{1 + \text{SNR} * G} \exp\left(-\frac{\text{SNR} |\mathbf{r}^\dagger \mathbf{M}_0^{-1} \mathbf{s}|^2}{1 + \text{SNR} * G}\right),$$

$$G = \mathbf{s}^\dagger \mathbf{M}_0^{-1} \mathbf{s}, \quad (22)$$

where $\mathbf{M}_0 = \sum_{k=1}^K \text{INR}_k \mathbf{f}_k \mathbf{f}_k^\dagger + \mathbf{I}_N$, SNR , G , \mathbf{s} , INR_k , and \mathbf{f}_k denote $\text{SNR}(\Psi, \mathbf{S})$, $G(\Psi, \mathbf{S})$, $\mathbf{s}(\Psi, \mathbf{S})$, $\text{INR}_k(\Psi, \mathbf{S}_k)$, and $\mathbf{f}_k(\Psi, \mathbf{S}_k)$, respectively, for simplicity.

Since SNR and G are constants under the assumption of a known environment, λ' , a monotonic function of the likelihood ratio, is also optimal:

$$\lambda'(\mathbf{r}) = |\mathbf{r}^\dagger \mathbf{M}_0^{-1} \mathbf{s}|^2. \quad (23)$$

Since λ' can be considered a chi-square random variable with two degrees of freedom for both the H_1 and H_0 hypotheses, an accurate analytical ROC expression for the optimal Bayesian detector is available using similar derivations as in Ref. 16. First, the probability density functions of λ' to both hypotheses are given by

$$p(\lambda' | H_0) = \exp(\lambda' / G) / G, \quad (24)$$

$$p(\lambda' | H_1) = \frac{1}{\text{SNR} * G^2 + G} \exp(\lambda' / (\text{SNR} * G^2 + G)).$$

From signal detection theory,

$$P_F(\beta) = \int_{\beta}^{\infty} d\lambda' p(\lambda' | H_0), \quad (25)$$

$$P_D(\beta) = \int_{\beta}^{\infty} d\lambda' p(\lambda' | H_1).$$

Substituting Eq. (24) in Eq. (25) yields

$$P_F(\beta) = \exp(-\beta / G), \quad (26)$$

$$P_D(\beta) = \exp(-\beta / (\text{SNR} * G^2 + G)).$$

Combining P_D and P_F results in the ROC expression for the Bayesian detector in the presence of interference in known environments.

$$P_D = P_F^{1/(\text{SNR} * G + 1)}, \quad G = \mathbf{s}^\dagger \mathbf{M}_0^{-1} \mathbf{s}. \quad (27)$$

If only a single interferer is present, substituting $\mathbf{M}_0^{-1} = \mathbf{I} - \text{INR} \mathbf{f} \mathbf{f}^\dagger / (1 + \text{INR})$ in Eqs. (22) and (27) yields the detector statistic

$$\lambda'(\mathbf{r}) = |\mathbf{r}^\dagger \left(\mathbf{I} - \frac{\text{INR}}{1 + \text{INR}} \mathbf{f} \mathbf{f}^\dagger \right) \mathbf{s}|^2, \quad (28)$$

and the performance prediction

$$P_D = P_F^{1/(\text{SNR} * G + 1)}, \quad G = \left(1 - \frac{\text{INR}}{1 + \text{INR}} |\rho|^2 \right), \quad (29)$$

where \mathbf{f} denotes the known interference wavefront, and $\rho = \mathbf{s}^\dagger \mathbf{f}$ is the correlation coefficient between the signal wavefront and the interference wavefront.

With the simple analytical ROC expressions, the effect of interference on detection performance is captured by the role of the key parameter $G = \mathbf{s}^\dagger \mathbf{M}_0^{-1} \mathbf{s}$. Since the covariance matrix \mathbf{M}_0 is positive definite, G is a positive number. Since the eigenvalues of \mathbf{M}_0 are equal or greater than one, the

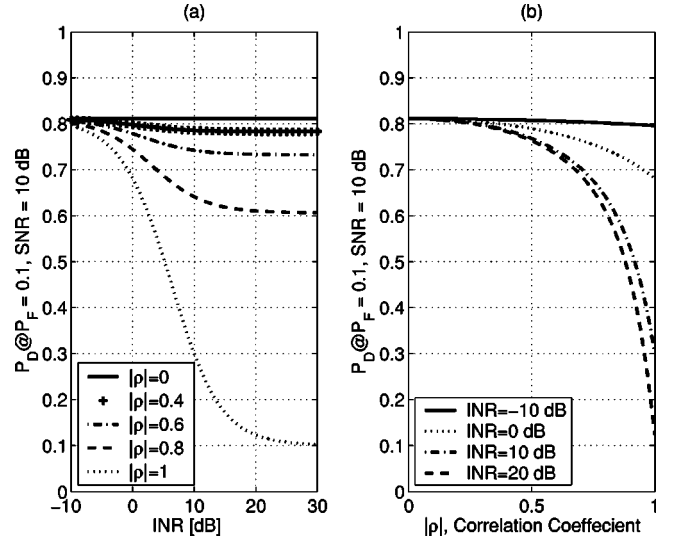


FIG. 1. Detection performance prediction in the presence of a single interference in a known ocean environment [Eq. (28)]. (a) P_D as a function of the interference-to-noise ratio, INR, at fixed $P_F = 0.1$ for various ρ values; (b) P_D as a function of ρ at fixed $P_F = 0.1$ for various INRs.

eigenvalues of \mathbf{M}_0^{-1} are in the range of zero and one. A singular value decomposition of \mathbf{M}_0^{-1} generates $\mathbf{M}_0^{-1} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\dagger$, where \mathbf{U} is a unitary matrix and $\mathbf{\Lambda}$ is a diagonal matrix composed of the eigenvalues of \mathbf{M}_0^{-1} . Since the eigenvalues are in the range of zero and one, $G = \mathbf{s}^\dagger \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\dagger \mathbf{s} \leq \mathbf{s}^\dagger \mathbf{U} \mathbf{U}^\dagger \mathbf{s} = \mathbf{s}^\dagger \mathbf{s} = 1$. Therefore, G is in the range of zero and one. Considering G is a weighting factor to the SNR in the ROC expression in Eq. (27), this range of G means that the presence of interference always decreases the SNR and consequently decreases the detection performance. The stronger the interference, the closer the G approaches zero, and the greater the performance degrades. On the other hand, if there is no interference, G is one, which does not impact the detection performance.

In the presence of a single interferer, G is determined by the INR and $|\rho|$, as shown in Eq. (29). The INR represents the energy of the interference as compared to the variance of diffuse noise at the receivers. The $|\rho|$ characterizes spatial similarity between the interference wavefront and the signal wavefront. This case is illustrated in Fig. 1. In Fig. 1(a), P_D is plotted as a function of INR for a set of $|\rho|$'s. It shows that the stronger the interference, the worse the detection performance. However, performance degradation is modest if the correlation coefficient between the interference and signal wavefronts is small. In Fig. 1(b), P_D is plotted as a function of $|\rho|$ for a set of INRs. It shows that when $|\rho|$ is less than 0.5, a single interferer has little impact on the detection performance, even if its energy is very strong.

The effect of interference on detection performance in known environments has been studied previously using different performance metrics.^{12,13} The performance metric used in Ref. 12 is a distance measure, and that used in Ref. 13 is array gain. Here we use the ROC metric, which is directly relevant to the ultimate goal of target detection. Figure 6.11 in Ref. 13 shows that the stronger the interference, the better the optimum array gain performance, which is different from our ROC performance result, as shown in Fig. 1.

The optimum array gain is an empirical metric defined as the array output to an input signal-to-noise ratio. When the noise field is diffuse noise only, the array gain metric is consistent with the ROC metric. In the presence of interference, the array gain metric cannot completely capture the effect of interference on detection performance. The utilization of array gain as the detection performance metric in the presence of interference can result in misleading conclusions.

VI. THE EFFECT OF UNDERESTIMATING THE NUMBER OF INTERFERERS

The Bayesian detector developed in this paper is designed for a known number of interferers. In practice, the Bayesian detection performance may be affected by the wrong *a priori* knowledge of the number of interferers. In this section, the effect of underestimating the number of interferers on Bayesian detection performance is quantified for a known ocean environment.

The data vector to both hypotheses are given by

$$\begin{aligned} H_1: \mathbf{r} &= \sqrt{\text{SNR}_i} a \mathbf{s}_i + \sqrt{\text{INR}_{K+1}} b_{K+1} \mathbf{f}_{K+1} + \mathbf{n}_1, \\ & i \in 1 \cdots M, \\ H_0: \mathbf{r} &= \sqrt{\text{INR}_{K+1}} b_{K+1} \mathbf{f}_{K+1} + \mathbf{n}_1, \\ \mathbf{n}_1 &= \sum_{k=1}^K \sqrt{\text{INR}_k} b_k \mathbf{f}_k + \mathbf{n}_0, \quad \mathbf{n}_0 \sim \mathcal{N}(0, \mathbf{I}_N). \end{aligned} \quad (30)$$

Equation (30) differs from Eq. (8) in the number of interferers.

Substituting Eq. (30) in Eq. (23) yields the test statistic

$$\begin{aligned} (\lambda'_{\text{mismatch}} | H_0) &= |(\text{INR}_{K+1} b_{K+1} \mathbf{f}_{K+1} + \mathbf{n}_1)^\dagger \mathbf{M}_0^{-1} \mathbf{s}|^2, \\ (\lambda'_{\text{mismatch}} | H_1) &= |(\sqrt{\text{SNR}_i} a \mathbf{s}_i + \text{INR}_{K+1} b_{K+1} \mathbf{f}_{K+1} \\ & \quad + \mathbf{n}_1)^\dagger \mathbf{M}_0^{-1} \mathbf{s}|^2. \end{aligned} \quad (31)$$

Let $x_0 = (\sqrt{\text{INR}_{K+1}} b_{K+1} \mathbf{f}_{K+1} + \mathbf{n}_1)^\dagger \mathbf{M}_0^{-1} \mathbf{s}$ and $x_1 = (\sqrt{\text{SNR}_i} a \mathbf{s}_i + \sqrt{\text{INR}_{K+1}} b_{K+1} \mathbf{f}_{K+1} + \mathbf{n}_1)^\dagger \mathbf{M}_0^{-1} \mathbf{s}$. The absolute squares of both x_0 and x_1 are chi-square random variables, with two degrees of freedom. Their probability density functions are

$$\begin{aligned} p(\lambda'_{\text{mismatch}} | H_0) &= \exp(\lambda' / D_0) / D_0, \\ p(\lambda'_{\text{mismatch}} | H_1) &= \exp(\lambda' / D_1) / D_1, \\ D_0 &= G + G_m, \quad D_1 = G + \text{SNR} * G^2 + G_m, \end{aligned} \quad (32)$$

where

$$G = \mathbf{s}^\dagger \mathbf{M}_0^{-1} \mathbf{s}, \quad G_m = \text{INR}_{K+1} |\mathbf{s}^\dagger \mathbf{M}_0^{-1} \mathbf{f}_{K+1}|. \quad (33)$$

Substituting Eqs. (32) and (33) in Eq. (25) and combining P_F and P_D yields the analytical ROC expression

$$P_{D_{\text{mismatched}}} = P_F^{1/(1 + \text{SNR} * G / G_m)}. \quad (34)$$

Equation (34) differs from Eq. (27) in that the effect of ignoring the $K+1$ th interferer on the Bayesian detection performance is captured by a degrading factor, G_m , to the SNR. From the G_m expression, we see that the higher the interference source level, and the higher the term $\mathbf{s}^\dagger \mathbf{M}_0^{-1} \mathbf{f}_{K+1}$, the

stronger degradation caused by underestimating the interferer number.

Substituting $K=0$ in Eq. (33) yields the degrading factor for the Bayesian detection algorithm that assumes diffuse noise only, while the true noise field is composed of a single interference plus diffuse noise:

$$G_m = 1 + \text{INR} |\rho|^2, \quad \rho = \mathbf{s}^\dagger \mathbf{f}. \quad (35)$$

VII. PERFORMANCE PREDICTIONS IN UNCERTAIN ENVIRONMENTS

Detection performance predictions become more difficult in the presence of interference and environmental uncertainty. Both factors can greatly limit the detection performance. Quantitative descriptions about the role of each factor on the detection performance are of practical interest. In this section, the analytical ROC expression is derived for the Bayesian detector in the presence of interference in an uncertain environment.

In Ref. 16, where the noise field is assumed to be diffuse noise only, the resultant analytical ROC expression for the Bayesian detector in uncertain environments is

$$P_D = 1 - (1 - P_F)^{(R-1)/R} (1 - (1 - (1 - P_F)^{1/R})^{1/(\text{SNR}+1)}), \quad (36)$$

where R is the rank of the signal matrix, characterizing the degree of environmental uncertainty. The SNR is the array-level signal-to-noise ratio at the receivers, assumed in the derivation to be constant over environmental uncertainty.

Here we transform the problem to the detection problem defined in Ref. 16 and exploit the result in Ref. 16. First, a prewhitening procedure is performed on the noise field. A singular value decomposition of the covariance matrix of the noise field generates $\hat{\mathbf{M}}_0 = \mathbf{U}_f \Lambda_f \mathbf{U}_f^\dagger$, where \mathbf{U}_f is a unitary matrix and Λ_f is a diagonal matrix composed of the eigenvalues of $\hat{\mathbf{M}}_0$. If $\mathbf{Q} = \Lambda_f^{-1/2} \mathbf{U}_f^\dagger$, then \mathbf{Q} is full rank. Since multiplying a full rank matrix on both sides of Eq. (8) does not change the detection performance, the detection problem becomes

$$\begin{aligned} H_1: \mathbf{r} &= \sqrt{\text{SNR}_i} a \mathbf{Q} \mathbf{s}_i + \mathbf{n}, \quad i \in 1 \cdots M, \\ H_0: \mathbf{r} &= \mathbf{n}, \quad \mathbf{n} \sim \mathcal{N}(0, \mathbf{I}_N), \end{aligned} \quad (37)$$

where the noise field \mathbf{n} is diffuse noise only.

Then the portion of the received signal, $\sqrt{\text{SNR}_i} a \mathbf{Q} \mathbf{s}_i$, is rewritten to fit into the framework of the detection problem defined in Ref. 16. The new signal wavefront vector is defined as

$$\begin{aligned} \mathbf{s}'_i &= \mathbf{Q} \mathbf{s}_i / \sqrt{(\mathbf{Q} \mathbf{s}_i)^\dagger \mathbf{Q} \mathbf{s}_i}, \\ &= \mathbf{Q} \mathbf{s}_i / \sqrt{\mathbf{s}_i^\dagger \hat{\mathbf{M}}_0^{-1} \mathbf{s}_i}, \\ &= \mathbf{Q} \mathbf{s}_i / \sqrt{G_i}. \end{aligned} \quad (38)$$

The new signal wavefront vector preserves the property $\mathbf{s}'_i{}^\dagger \mathbf{s}'_i = 1$.

Substituting Eq. (38) in Eq. (37) results in

$$\begin{aligned} H_1: \mathbf{r} &= \sqrt{\text{SNR}_i G_i} a \mathbf{s}'_i + \mathbf{n}, \quad i \in 1 \cdots M, \\ H_0: \mathbf{r} &= \mathbf{n}, \quad \mathbf{n} \sim \mathcal{N}(0, \mathbf{I}_N). \end{aligned} \quad (39)$$

Now the detection problem in the presence of interference is translated to the detection problem in diffuse noise,¹⁶ with the product of the SNR_i and G_i replacing SNR_i , and with the new signal matrix $\mathfrak{R}' = [s'_1, s'_2, \dots, s'_M]$ replacing the original signal matrix \mathfrak{R} . The SNR_i is the signal-to-noise ratio at the receivers, which is assumed constant in the derivations in Ref. 16. The parameter G_i is the signal-to-interference coefficient. An additional assumption used here is that G_i is constant over environmental uncertainty, i.e., $G_i = G$ for $i = 1 \dots M$. Since $\mathfrak{R}' = \mathbf{Q}\mathfrak{R}/\sqrt{G}$ and the transformation \mathbf{Q}/\sqrt{G} is full rank, the rank of the new signal matrix is equal to the rank of the original signal matrix, which is R .

Using similar steps as in Ref. 16, an analytical ROC expression for the Bayesian detector in the presence of interference in uncertain environments is obtained. This ROC expression differs from Eq. (36) in that the product of the SNR and G replaces the SNR:

$$P_D = 1 - (1 - P_F)^{(R-1)/R} \times (1 - (1 - (1 - P_F)^{1/R})^{1/(\text{SNR} * G + 1)}). \quad (40)$$

In real applications, the SNR and G could be replaced by the average SNR and the average G over realizations of the uncertain environmental parameters, i.e.,

$$\text{SNR} = \frac{1}{M} \sum_{i=1}^M \text{SNR}_i, \quad G = \frac{1}{M} \sum_{i=1}^M G_i. \quad (41)$$

For a short vertical array or horizontal array, the constant SNR and constant G approximation might fail because the received signal energy is sensitive to source range variation. The application of Eq. (40) for an optimal Bayesian performance prediction is limited to scenarios,¹⁶ where the received signal energy, the received interference energy, and the noise level are relatively stable.

Equation (40) is the key result in this paper. This simple analytical ROC expression captures the effects of the presence of both interference and environmental uncertainty on detection performance. The effect of the presence of interference is represented by the role of the parameter G . In Sec. V, we have shown that G is in the range of zero and one for a known environment. It can be shown that G is in the range of zero and one for an uncertain environment. This means that the presence of interferers always decreases the detection performance, in either a known or uncertain environment. However, in uncertain environments, the computation of G is affected by the presence of environmental uncertainty, through the covariance matrix $\hat{\mathbf{M}}_0$.

Figure 2 illustrates Eq. (40) by plotting P_D as a function of the key parameters: the signal-to-interference coefficient G and the environmental uncertainty scale R . Figure 2(a) plots P_D as a function of $1 - G$, assuming $P_F = 0.1$, $\text{SNR} = 10$ dB, for $R = 1, 10$, and 100 . It shows that for $R = 1$, i.e., a known environment, P_D starts at 0.81 for $G = 1$, then quickly drops to 0.1 for $G = 0$. While for $R = 100$, i.e., a greatly uncertain environment, P_D drops from 0.59 to 0.1 for the same range of G . Figure 2(a) indicates that when G is close to zero, the detection performance is primarily determined by G rather than R ; when G is close to one, the detection performance is independently determined by G and

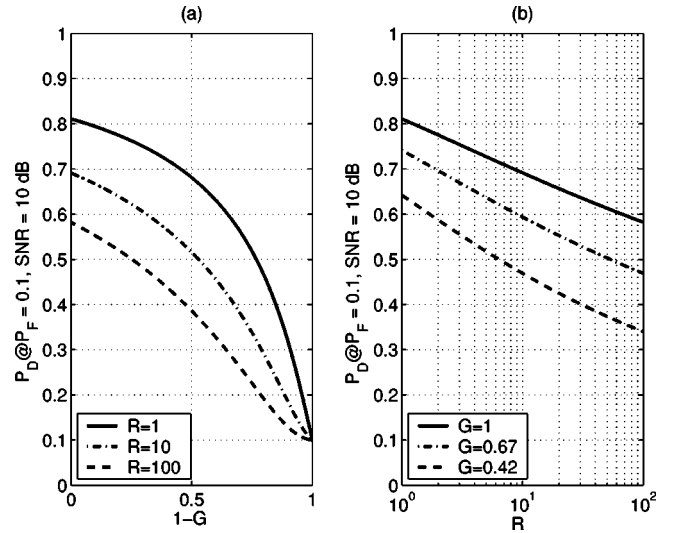


FIG. 2. Detection performance prediction in the presence of interference in an uncertain ocean environment [Eq. (36)]. (a) P_D as a function of $1 - G$ at fixed $P_F = 0.1$, $\text{SNR} = 10$ dB for various R 's, where G is the signal-to-interference coefficient and R is the rank of the signal matrix; (b) P_D as a function of R at fixed $P_F = 0.1$, $\text{SNR} = 10$ dB for various G 's.

R . Figure 2(b) plots P_D as a function of R on a logarithmic scale, assuming $P_F = 0.1$ and for $G = 1, 0.67$, and 0.42 . It shows that the performance degradation is inversely proportional to R on the logarithmic scale for all three G conditions.

Equation (40) is a general result that is consistent with the results for special cases. For example, substituting $R = 1$ in Eq. (40) yields the ROC performance prediction expression for known environments and substituting $G = 1$ in Eq. (40) generates the results for diffuse noise only.¹⁶ Generality is an advantage of this detection performance prediction approach in addition to simplicity and computational feasibility. The developed detector statistics and the ROC performance predictions along with the expressions for special cases are summarized in Table I.

VIII. SIMULATION RESULTS

Simulation scenarios extended from the general mismatch benchmark shallow-water model ("genlmis") proposed in the May 1993 NRL Workshop on Acoustic Models in Signal Processing^{17,5,8} were used to study the effects of interference on detection performance and to check the approximate analytical ROC expressions. Figure 3 reviews a modified "genlmis" model. This model also consists of a single signal source, a single interferer, and a vertical line receiving array contains 100 hydrophones spaced 1 m apart, ranging from 1 to 100 m in depth. Table II summarizes the environmental parameters shown in Fig. 3. Three environmental configuration scenarios are used. The known ocean scenario means that the environmental parameters are configured with their mean values and are known *a priori*. The uncertain water depth scenario assumes that the water depth parameter is uncertain as defined in Table II, and other parameters take the mean values. The general uncertain scenario considers seven uncertain parameters defined in Table II: D , C_0 , C_D^- , C_D^+ , C_l , ρ , and α .

TABLE I. Taxonomy of results for sonar detection performance prediction.

Problem	Detector statistic	ROC expression
Known ocean		
Interference + diffuse noise	$\lambda'(\mathbf{r}) = \mathbf{r}^\dagger \mathbf{M}_0^{-1} \mathbf{s} ^2$ Eq. (23)	$P_D = P_F^{1/(\text{SNRs}^\dagger \mathbf{M}_0^{-1} \mathbf{s} + 1)}$ Eq. (27)
Single interference + diffuse noise	$\lambda'(\mathbf{r}) = \left \mathbf{r}^\dagger \left(\mathbf{I} - \frac{\text{INR}}{1 + \text{INR}} \mathbf{f} \mathbf{f}^\dagger \right) \mathbf{s} \right ^2$ Eq. (28)	$P_D = P_F^{1/(\text{SNR}(1 - \text{INR}/1 + \text{INR} \mathbf{s}^\dagger \mathbf{f} ^2) + 1)}$ Eq. (29)
Diffuse noise only	$\lambda'(\mathbf{r}) = \mathbf{r}^\dagger \mathbf{s} ^2$ Ref. 16	$P_D = P_F^{1/(\text{SNR} + 1)}$ Ref. 16
Uncertain ocean		
Interference + diffuse noise	$\lambda(\mathbf{r}) = \frac{1}{M} \sum_{i=1}^M \lambda_i,$ $\lambda_i = \frac{1}{1 + \text{SNR}_i G_i}$ $\times \exp\left(-\frac{\text{SNR}_i \mathbf{r}^\dagger \hat{\mathbf{M}}_0^{-1} \mathbf{s}_i ^2}{1 + \text{SNR}_i G_i}\right)$ Eq. (20)	$P_D = 1 - (1 - P_F)^{R-1/R}$ $\times (1 - (1 - (1 - P_F)^{1/R})^{1/\text{SNR}^* G + 1})$ Eq. (40)
Diffuse noise only	$\lambda(\mathbf{r}) = \frac{1}{M} \sum_{i=1}^M \lambda_i,$ $\lambda_i = \frac{1}{1 + \text{SNR}_i}$ $\times \exp\left(-\frac{\text{SNR}_i \mathbf{r}^\dagger \mathbf{s}_i ^2}{1 + \text{SNR}_i}\right)$ Ref. 16	$P_D = 1 - (1 - P_F)^{R-1/R}$ $\times (1 - (1 - (1 - P_F)^{1/R})^{1/\text{SNR} + 1})$ Ref. 16

The acoustic pressure field generated by a point source at depth z_s and range r_s is computed using the normal mode theory,²²

$$p(z; r_s, z_s) = p_0 \sum_{m=1}^K \Phi_m(z_s) \Phi_m(z_i) \times \exp(-jk_m r_s) / \sqrt{k_m r_s}, \quad (42)$$

where z is the depth of the receiver and p_0 is a normalization constant. The number of propagating acoustic modes is represented by K . The terms k_m and Φ_m are the m th eigenvalue and eigenfunction of the Sturm–Liouville problem, which can be calculated using the KRAKEN²³ code.

Figure 4 illustrates the effect of the presence of a single interferer on the Bayesian detection performance using the known ocean scenario. The signal source is fixed at 5950 m range and 55 m depth. The interferer range and depth are

known *a priori*. The P_D performance is computed using Eq. (29). In Fig. 4, P_D at fixed $P_F = 0.1$ is plotted as a function of the range and depth of the interferer, for about 10 dB SNR at the receivers. In Figs. 4(a) and 4(b), the interferer levels are the same as, or 30 dB more than, the signal source level. The plots show that in the known benchmark ocean, the presence of a single interferer does not impact the detection performance, even though the interferer level is very high. This is determined by the spatial correlation property of the benchmark ocean. In the known benchmark ocean, the absolute value of the correlation coefficient between the wavefronts that come from two arbitrarily separated sources is very weak ($|\rho| < 0.2$), if the source separation is greater than 30 m in range and 4 m in depth.

The effect of mismatched *a priori* interferer number on the Bayesian detection performance is illustrated in Fig. 5 using the known ocean scenario. The scenario includes a

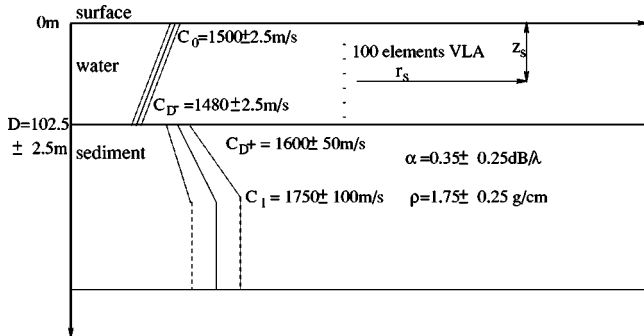


FIG. 3. NRL shallow-water environmental model.

TABLE II. Parameters of uncertain shallow-water environments.

Environmental parameter	Symbol	Value Range
Bottom depth	D	102.5 ± 2.5 m
Surface sound speed	C_0	1500 ± 2.5 m/s
Bottom sound speed	C_D^-	1480 ± 2.5 m/s
Upper sediment sound speed	C_D^+	1600 ± 50 m/s
Lower sediment sound speed	C_I	1750 ± 100 m/s
Sediment density	ρ	1.7 ± 0.25 g/cm ³
Sediment attenuation	α	0.35 ± 0.25 dB/λ
Sediment thickness	t	100 m

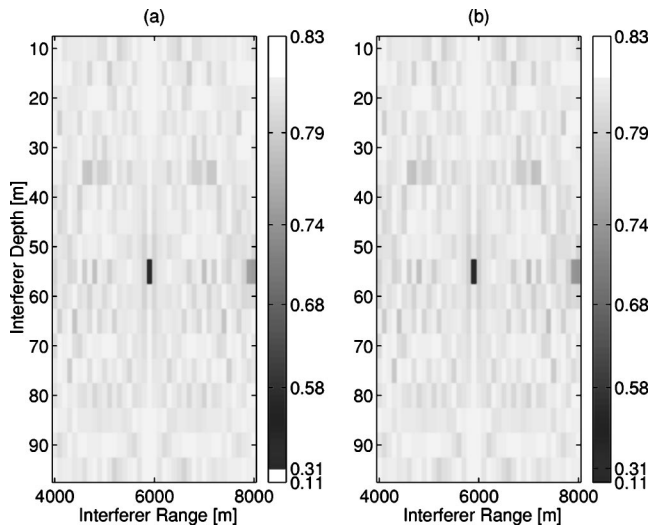


FIG. 4. Detection performance in the presence of a single interferer in the known benchmark ocean. The P_D image is plotted at fixed $P_F=0.1$ for SNR=10 dB at receivers. The coordinates of the image are the range and depth of the interferer. The signal source is located at 5950 m in range and 55 m in depth. (a) $\sigma_b^2 = \sigma_a^2$; (b) $\sigma_b^2 = 1000\sigma_a^2$.

single signal source fixed at 6000 m range and 50 m depth, and a single interferer fixed at 7 km range and 10 m depth. In Fig. 5, the predicted P_D performance is plotted as a function of the INR for three SNR conditions. Solid curves are P_D computed using Eq. (29), where the detector uses the correct *a priori* information about the interferer. Dashed curves are for P_D computed using Eq. (35), where the detector assumes diffuse noise only. The plots show that the degradation of the detection performance can be significant if the INR of the ignored interferer is comparable to the SNR.

The approximate analytical ROC expression for the Bayesian detector in uncertain environments is verified by comparing its prediction results with those obtained with Monte Carlo evaluations. Both the analytical approach and the Monte Carlo approach assume that the ranges of environ-

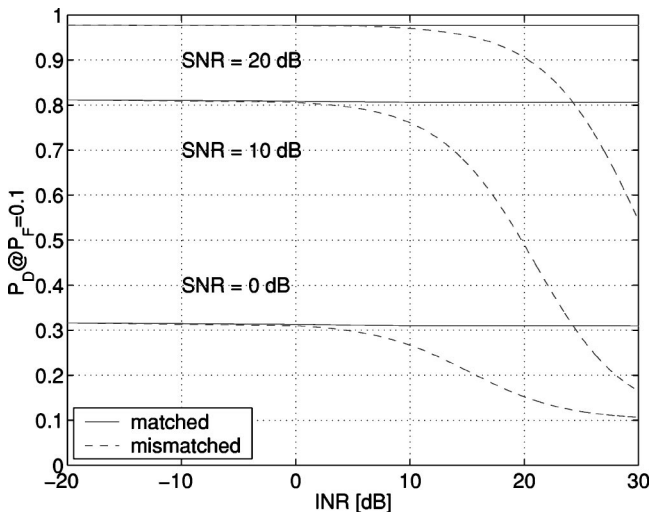


FIG. 5. Degradation of Bayesian detection performance due to underestimating the number of interferers in the known benchmark ocean. The P_D is plotted as a function of INR at fixed $P_F=0.1$ for SNR=20, 10, and 0 dB, for the Bayesian detector that uses correct (solid curve) or wrong (dashed curve) *a priori* knowledge about the interferer number.

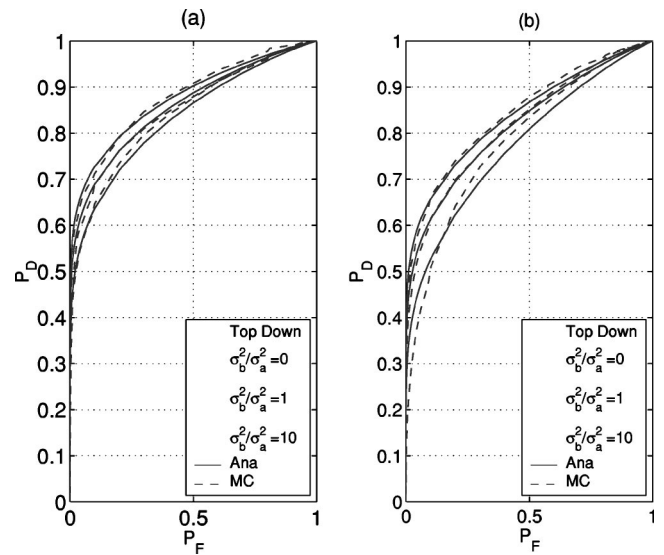


FIG. 6. The analytical ROC performance predictions in the presence of an interferer with various INR values in uncertain environments [Eq. (36)], comparing with Monte Carlo performance evaluations of the Bayesian Detector [Eq. (21)]. (a) Uncertain ocean depth scenario; (b) general uncertain scenario.

mental uncertainties, the signal variance to diffuse noise variance ratio σ_a^2/σ_n^2 , and the interference variance to diffuse noise variance ratio σ_b^2/σ_n^2 were known *a priori*. The simulated data was generated using Monte Carlo sampling techniques over different ocean environmental, source position, diffuse noise, signal amplitude and phase, interference amplitude and phase realizations based on the data model [Eq. (8)]. In using the Monte Carlo approach, 5000 data samples were generated for each hypothesis to produce the samples of the likelihood ratio [Eq. (21)] and to compute P_F and P_D [Eq. (25)]. Before generating the likelihood ratio samples, the signal matrix and the interference matrix were generated using 500 realizations of the signal wavefronts ($M=500$, $L=500$). The SNR_{*i*}, \hat{M}_0 , and INR_{*i*} were computed using Eqs. (2), (14), and (4), respectively. In using the analytical approach, three parameters: the signal-to-noise ratio, the signal-to-interference coefficient, and the rank of the signal matrix were estimated. The rank of the signal matrix was estimated by counting the number of significant eigenvalues of the signal matrix that exceed a threshold: 5% of the maximum eigenvalue. The signal-to-interference coefficients were computed for each of the SNR and interferer level conditions using Eqs. (21), (14), and (41). The SNR was computed using Eqs. (2) and (41).

The primary cost for the Monte Carlo approach is due to generating samples of the likelihood ratio for each hypothesis, which is about $O(MNP)$, where P is the number of samples, M is the number of columns in the signal matrix, N is the number of array sensors. The primary cost for the analytical approach is due to the computation of the eigenvalues in order to estimate the rank of the signal matrix and the computation of the matrix inverse to estimate G , which is $O(N^3)$.^{24,25} Typically, M is greater than N , and P is above 1000 so that the precision of P_D and P_F reaches 0.001. P is about 10 times greater than N . Therefore, the analytical approach speeds up more than an order of magnitude.

TABLE III. Estimated signal-to-interference coefficient.

Scenario	Interferer level	SNR			
		0 dB	10 dB	20 dB	30 dB
Known	$\sigma_b^2/\sigma_a^2=1$	0.9859	0.9700	0.9662	0.9658
Environment	$\sigma_b^2/\sigma_a^2=10$	0.9700	0.9662	0.9658	0.9657
Uncertain	$\sigma_b^2/\sigma_a^2=1$	0.9667	0.8470	0.6927	0.5969
Water depth	$\sigma_b^2/\sigma_a^2=10$	0.8470	0.6927	0.5969	0.5398
General	$\sigma_b^2/\sigma_a^2=1$	0.9683	0.8316	0.5977	0.4017
Uncertain	$\sigma_b^2/\sigma_a^2=10$	0.8316	0.5977	0.4017	0.2428

Figure 6 illustrates the analytical and the Monte Carlo performance prediction results by plotting ROC curves for three interferer levels, at SNR=10 dB, using the uncertain water depth scenario and the general uncertain scenario. The plots show that the ROCs generated using the approximate analytical approach agree with those obtained using the Monte Carlo evaluation approach for both scenarios, and for various interferer levels. The results indicate that the proposed Bayesian performance prediction ROC expression captures the primary effects of the presence of interference and environmental uncertainty by the roles of the two fundamental parameters: the signal-to-interference coefficient and the rank of the signal matrix.

The effect of interference on Bayesian detection performance in uncertain environments is represented by the signal-to-interference coefficient. Table III lists the signal-to-interference coefficient values for three scenarios and various SNRs and interferer levels. It shows that G depends on the scale of environmental uncertainty. For example, for the entries of SNR=10 dB and $\sigma_b^2/\sigma_a^2=10$ (the interferer level is 10 times stronger than the target source level), G decreases from 0.9662 for the known environment scenario to 0.6927 for the uncertain water depth scenario and then to 0.5977 for the general uncertain scenario. The degrees of environmental uncertainty, i.e., the rank of the signal matrix, are 1, 5, and 13, respectively, for these three scenarios. The result shows that G decreases with increased environmental uncertainty, which means that the degradation effect of the interference on detection performance is strengthened by environmental uncertainty. This is because the rank of the interference matrix also increases with environmental uncertainty, resulting in an increase in the number of effective eigenvalues that inversely impact the computation of G . The higher the interferer level, the higher the eigenvalues for the covariance matrix, and the greater the inverse impact on the G values.

Figure 7 illustrates the combined effect of interference and environmental uncertainty on the detection performance by plotting P_D as a function of the SNR for a fixed $P_F=0.1$ for three interference levels: $\sigma_b^2/\sigma_a^2=0, 1, 10$. The uncertain water depth scenario and the general uncertain scenario were used. In each plot, the benchmark P_D performance predictions, denoted by solid lines, are computed for the known ocean scenario. Since the absolute value of the correlation coefficient between the known signal and interference wavefronts is very small, three benchmark P_D curves for three interference levels are indistinguishable. The dashed curves are analytical prediction results [Eq. (40)] and the circles are Monte Carlo evaluation results for the Baye-

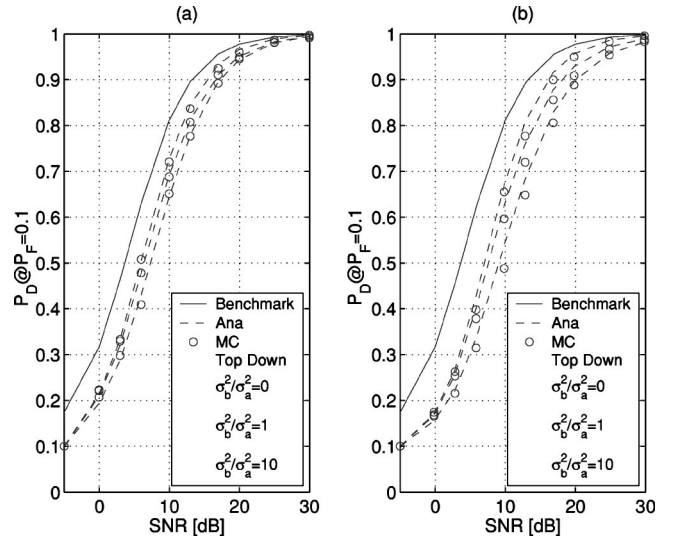


FIG. 7. Detection performance prediction P_D as a function of SNR in the presence of an interferer in uncertain environments [Eq. (36)], comparing with Monte Carlo performance evaluations of the Bayesian detector [Eq. (21)]. (a) Uncertain ocean depth scenario; (b) general uncertain scenario.

sian detector [Eq. (21)]. The degradation of the detection performance due to interference and environmental uncertainty can be illustrated by an increased SNR threshold in order to achieve a fixed P_D . For example, to achieve a fixed $P_D=0.8$, Fig. 7(a) shows that the SNR threshold is 10 dB for the known environment scenario and are about 12, 13, and 14 dB for the uncertain water depth scenario ($R=5$) for three increased interferer levels. The uncertainty of the uncertain water scenario results in a 2–4 dB SNR increase. Figure 7(b) shows that the uncertainty of the general uncertain scenario ($R=13$) results in a 2.5–7 dB SNR increase. Thus, detection performance degradation due to interference is magnified by the scale of environmental uncertainty. On the other hand, the results also indicate that the detection performance degrades much faster with an increased degree of environmental uncertainty in the presence of interference as compared to that for diffuse noise only scenarios.¹⁶

IX. CONCLUSIONS

Analytical ROC performance prediction expressions, which are computationally fast, are developed for the Bayesian detector in the presence of interference in uncertain environments. This is a significant extension to the previous detection performance prediction approach proposed for a diffuse noise only circumstance.¹⁶ The ROC expression is developed within a Bayesian decision framework that incorporates uncertainties in ocean environmental parameters and source (target and interferer) position parameters.

The analytical ROC expression is verified using several uncertain environment scenarios extended from benchmark propagation models. The results demonstrate that the ROC expression is a good approximation to the ROC obtained using Monte Carlo techniques, and can be computed much faster. The simple ROC expression captures fundamental parameters that impact sonar detection performance: the signal-to-noise ratio, the rank of the signal matrix, and the signal-to-interference coefficient. It provides a more realistic

performance prediction than the classic sonar equation that fails to incorporate the effects of interference and environmental uncertainty.

The signal-to-interference coefficient in the ROC expression characterizes the effect of interference on detection performance in uncertain environments. In computing the signal-to-interference coefficient, environmental uncertainty is incorporated through the covariance matrix of the noise field, which is obtained based on the *a priori* information of uncertain environmental parameters and uncertain interferer position parameters. The values of the signal-to-interference coefficient in different simulation scenarios show that environmental uncertainty magnifies the degrading effect of interference on the detection performance.

ACKNOWLEDGMENT

This work was supported by the ONR Uncertainty Defense Research Initiative (DRI).

- ¹F.-P. A. Lam, F. P. A. Benders, P. Schippers, and S. P. Beerens, "Environment adaptation for LFAS and corresponding 3D modeling of sonar performance," "UDT Europe 2002" Undersea Defence Technology Conference and Exhibition, La Spezia, Italy, 18–20 June, 2002.
- ²A. Tolstoy, "Sensitivity of matched-field processing to sound-speed profile mismatch for vertical arrays in a deep-water pacific environment," *J. Acoust. Soc. Am.* **85**, 2394–2404 (1989).
- ³H. Schmidt, A. B. Baggeroer, W. A. Kuperman, and E. K. Sheer, "Environmentally tolerant beamforming for high-resolution matched field processing: Deterministic mismatch," *J. Acoust. Soc. Am.* **88**, 1802–1810 (1990).
- ⁴A. M. Richardson and L. W. Nolte, "A posteriori probability source localization in an uncertain sound speed, deep ocean environment," *J. Acoust. Soc. Am.* **89**, 2280–2284 (1991).
- ⁵J. A. Shorey, L. W. Nolte, and J. L. Krolik, "Computationally efficient Monte Carlo estimation algorithm for matched field processing in uncertain ocean environments," *J. Comput. Acoust.* **2**, 285–314 (1994).
- ⁶J. L. Krolik, "Matched field minimum variance beamforming in a random ocean channel," *J. Acoust. Soc. Am.* **92**, 1408–1419 (1992).
- ⁷N. Lee, L. M. Zurk, and J. Ward, "Evaluation of reduced-rank, adaptive matched field processing algorithms for passive sonar detection in a shallow-water environment," *Signals, Systems, and Computers, 1999*, Conference Record of the 33rd Asilomar Conference 1999, Vol. 2, pp. 876–880.
- ⁸A. Ephraty, J. Tabrikian, and H. Messer, "Underwater source detection using a spatial stationarity test," *J. Acoust. Soc. Am.* **109**, 1053–1063 (2001).
- ⁹V. Premus, D. Alexandrou, and L. W. Nolte, "Full-field optimum detection in an uncertain, anisotropic random wave scattering environment," *J. Acoust. Soc. Am.* **98**, 1097 (1995).
- ¹⁰J. L. Krolik, "The performance of matched-field beamformers with Mediterranean vertical array data," *IEEE Trans. Signal Process.* **44**, 2605–2611 (1996).
- ¹¹M. Wazenski and D. Alexandrou, "Active, wideband detection and localization in an uncertain multipath environment," *J. Acoust. Soc. Am.* **101**, 1961 (1997).
- ¹²H. Messer, "The use of spectral information in optimal detection of a source in the presence of a directional interference," *IEEE J. Ocean. Eng.* **19**, 422–430 (1994).
- ¹³H. L. Van Trees, *Optimum Array Processing. Part IV of Detection, Estimation, and Modulation Theory* (Wiley Interscience, New York, 2002).
- ¹⁴I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.* **AES-10**, 853–863 (1974).
- ¹⁵F. C. Robby, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.* **28**, 208–216 (1992).
- ¹⁶L. Sha and L. W. Nolte, "Effects of environmental uncertainties on sonar detection performance prediction," *J. Acoust. Soc. Am.* (in press).
- ¹⁷M. B. Porter and A. Tolstoy, "The matched-field processing benchmark problem," *J. Comput. Acoust.* **2**, 161–185 (1994).
- ¹⁸K. Harmanci, J. Tabrikian, and J. L. Krolik, "Relationship between adaptive minimum variance beamforming and optimal source localization," *IEEE Trans. Signal Process.* **48**, 1–12 (2000).
- ¹⁹A. B. Baggeroer, "A 'true' maximum likelihood method for estimating frequency-wavenumber spectra," *Proc., Inst. Acoust.* **13**, 217–224 (1991).
- ²⁰W. W. Peterson, T. G. Birdsall, and W. C. Fox, "The theory of signal detectability," *Information Theory, IEEE Trans.* **4**(4), 171–212 (1954).
- ²¹D. H. Johnson and D. E. Dudgeon, *Array Signal Processing Concepts and Techniques* (Prentice-Hall Signal Processing, Englewood Cliffs, NJ, 1993).
- ²²F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics*, AIP Series in Modern Acoustics and Signal Processing (AIP Press, Woodbury NY, 1994).
- ²³M. B. Porter, "The KRAKEN normal mode program," SACLANT Undersea Research Centre, La Spezia, Italy, 1991.
- ²⁴Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, H. van der Vorst, and editors, *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide* (SIAM, Philadelphia, 2000).
- ²⁵T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms* (McGraw-Hill Book Company and the MIT Press, New York, 1998).

Concurrent inversion of geo- and bio-acoustic parameters from transmission loss measurements in the Yellow Sea

Orest Diachok^{a)} and Stephen Wales
Naval Research Laboratory, Washington, DC 20375

(Received 9 March 2003; revised 15 October 2004; accepted 25 October 2004)

This paper describes results of a simultaneous inversion of bio-acoustic parameters of fish (anchovies) and geo-acoustic parameters of the bottom from transmission loss (TL) measurements in the Yellow Sea, which were reported by Qiu *et al.* [*J. Sound Vib.* **220**, 331–342 (1999)]. This data set was selected because the bio-absorptivity at their site was extremely large, 40 dB at 1.3 kHz at 5 km, and measurements were made between multiple source and receiver depths and ranges. Measurements were made at night when anchovies are generally dispersed. Replica fields were calculated with a normal mode model, which incorporates bio-absorption layers. The inversion was based on minimizing the rms difference, Δ , between measured and calculated values of TL at all ranges and source and receiver depths, and involved a simultaneous search for bio-layer depth, bio-layer thickness, bio-alpha, geo-sound speed, and geo-alpha. The resultant small value of Δ , ± 1.7 dB, confirmed that the model, which was assumed in replica field calculations, was realistic, and that inverted parameters were meaningful. In particular, the inverted depth of the bio-absorption layer, 6.9 ± 0.3 m, was consistent with theoretical calculations of the depth, 5.8 ± 1 m, of 10-cm-long anchovies *Engraulis japonicus*, the dominant species in the Yellow Sea. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1862093]

PACS numbers: 43.30.Sf, 43.30.Ft [KGF]

Pages: 1965–1976

I. INTRODUCTION

Bio-acoustic absorptivity can have large (nominally as large as 40 dB at 5 km) frequency selective effects on transmission loss frequencies between approximately 0.5 and 5 kHz in shallow waters. In the mid-1960s Weston (1967) demonstrated that large changes in transmission loss occurred near sunrise and sunset over limited frequency bands. His experiments were conducted with several fixed sources (which operated between 0.4 and 4.4 kHz) and several fixed hydrophones on the bottom at a relatively shallow (~ 40 m) site in the Bristol Channel at ranges of 18, 22, and 120 km. Fixing the source and receiver eliminated changes in TL due to changes in the geo-acoustic properties of the bottom, minimized changes in TL versus time due to changes in oceanic sound speed, and permitted monitoring temporal changes in the magnitude of frequency selective attenuation due to fish. The latter was generally much higher at night, when fish such as sardines are dispersed near the surface. In this mode their separations are comparable to or greater than the wavelength at the resonance frequency of their swim bladders. The peak frequencies of losses at night, which occurred between 1 and 4.5 kHz, were consistent with the dimensions and depths of the swim bladders of sardines (Ching and Weston, 1971). The sardine was the dominant species in the Bristol Channel in the 1960s. However, Weston did not conduct concurrent trawling measurements of fish length distributions or echo-sounder measurements of fish layer depths. Consequently, he could not offer experimental proof of his hypothesis.

Subsequently Diachok conducted experiments, which included the following: (1) measurements of transmission loss as a function of time and frequency, (2) echo-sounder based measurements of the depths of bio-acoustic absorption layers, and (3) trawling-based measurements to determine the length distributions of the dominant species of swim-bladder-bearing fish. His experiment, which was conducted in the Gulf of Lion (Diachok, 1999), demonstrated that large, frequency-selective absorption losses occurred at the resonance frequencies of sardine and anchovy swim bladders. This experiment was conducted with a fixed broadband source and a fixed vertical receiving array. This geometrical configuration permitted inference of (1) the temporal changes in the frequencies of absorption lines (which were due to changes in the resonance frequencies of swim bladders as fish undergo vertical migrations at sunrise and sunset), (2) the average depth and thickness of absorption layers, and (3) the number densities (number per cubic meter) of fish within absorption layers. These parameters were derived by matching measured and calculated transmission loss as a function of depth and frequency. Resultant inferences were consistent with concurrent echo sounder and trawling measurements of these parameters.

In addition to absorption lines, which were associated with dispersed adult fish, these experiments also revealed the presence of absorption lines at lower and higher frequencies. A low-frequency line at a frequency equal to approximately $0.6f_0$ (where f_0 is the resonance frequency of individual sardines) during daytime in the Gulf of Lion was attributed to “bubble cloud” effects of adult sardine schools, in which average separations are approximately one fish length (Diachok, 1999). Such separations are much smaller than the wavelength at the resonance frequency. High-frequency lines

^{a)}Present address: Poseidon Sound, Oakton, VA 22124; electronic mail: OrestDia@aol.com

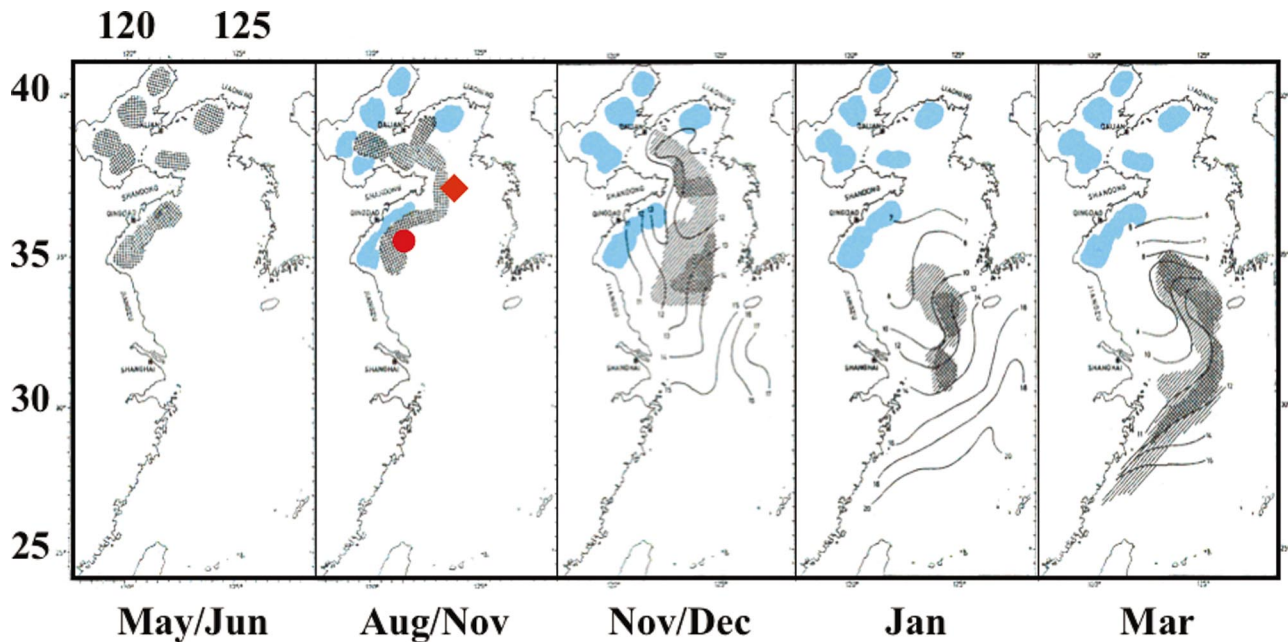


FIG. 1. Locations of transmission loss measurements by Qiu *et al.* (1999), the red dot (●), Dahl *et al.* (1998) and Rogers *et al.* (2000), the red diamond (◆) and “idealized” distributions of juvenile (blue) and adult (gray) anchovies in the Yellow Sea, based Deng and Zhao (1991) and Iverson *et al.* (1993). Adults spawn in May/June and migrate south in winter. Adults spawn *predominantly* near shore; juveniles *tend to remain* near shore.

were attributed to juvenile fish. An absorption line at 3.6 kHz in the Gulf of Lion was consistent with the presence of 6.5 cm long juvenile sardines (Diachok, 1999).

It is noteworthy that the peak frequencies of absorption lines (Diachok *et al.*, 2004) and back-scattering lines (Holliday, 1972) due to dispersed adult anchovies in the seas off California were essentially equal (at the same depth). This experimental observation is consistent with recently reported concurrent measurements of absorption and back-scattering spectra in the Yellow Sea (Qiu *et al.*, 1999), which revealed that the peak frequencies of concurrently measured absorption and back-scattering spectra at this site were nearly identical. Measured peak frequencies were consistent with calculations of the resonance frequency of swim bladders of anchovies, based on historical measurements of the average lengths and depths of this species in the Yellow Sea.

It is well known that transmission loss in shallow water is controlled by sound speed profiles in the water and the geo-acoustic parameters of the bottom. Experiments conducted by Ching and Weston (1971), Diachok, (1999), and Qiu *et al.* (1999) indicate that bio-acoustic absorptivity due to fish with swim bladders can also have a large effect on transmission loss. The objectives of this paper are to present the results of a concurrent inversion of the bio-acoustic parameters of fish (anchovies) and the geo-acoustic parameters of the bottom from Qiu *et al.*'s transmission loss data, and to demonstrate that inverted parameters are consistent with previously published values of these parameters in the Yellow Sea.

Sections II and III of this paper provide a review of the bio-acoustic properties and temporal and spatial variability of anchovies, the species of fish which is the most numerous and has the most dramatic effect on transmission loss in the Yellow Sea. Section IV provides a brief review of the geo-acoustic properties of Qiu *et al.*'s measurement site. BIO-C-

SNAP, a transmission loss model that permits simulations of the effects of bio-acoustic absorption layers, due to fish with swim bladders, and the effects of the bottom on transmission loss will be reviewed in Sec. V. This will be followed by a summary of Qiu *et al.*'s experimental procedures and their transmission loss data, and a discussion of the frequency and the Q of an apparent bio-absorption line in Sec. VI. Section VII describes the results of the concurrent inversion of the geo-acoustic parameters of the bottom and the bio-acoustic parameters of dispersed anchovies. Sections VIII and IX consist of discussions of inverted values of geo-alpha and bio-acoustic parameters, respectively. Section X provides a discussion of the uncertainty of inverted values, followed by conclusions in Sec. XI.

II. TEMPORAL AND SPATIAL DISTRIBUTIONS OF ANCHOVIES

Anchovies are the dominant fish in the Yellow Sea. According to Ohshimo (1996), about 80% of the fish, which he trawled in the northern Yellow Sea, were anchovies. Trawling is length selective: it discriminates against “small” fish, due to mesh size, and “large” fish, because they swim fast and are able to avoid capture. Consequently, his results represent an upper bound on the percentage of anchovies in this region. Anchovies are at present (Ohshimo, personal communication) and were the dominant species in this region throughout the 1990s (Ohshimo, 1996; Iverson *et al.*, 1993). Anchovies spawn near the coast of China, predominantly at sites shown in Fig. 1 (Iverson *et al.*, 1993; Deng and Zhao, 1991). Following spawning, juveniles tend to remain in the vicinity of spawning sites, whereas adults migrate south for the winter and return to their spawning sites in the spring, as

illustrated in Fig. 1. This figure should be considered an “idealized” version of this process. A significant percentage of spawnings also occurs throughout the Yellow Sea. Furthermore, currents and eddies can transport juveniles from the predominant spawning sites into the central regions of the Yellow Sea during their first year (Iverson *et al.*, 1993). Adult anchovies in this region are typically 10 ± 1 cm long (standard length) and their resonance frequencies at a depth of 7 m are nominally 1.4 kHz. The lengths of juveniles are smaller, and their resonance frequencies higher, being about 3.5 kHz at 7 m in their fourth month (Diachok, 2001). The second most common species in the Yellow Sea is jack mackerel (Deng and Zhao, 1991). These fish are typically about 20 cm long and resonate at about 0.6 kHz at 7 m (Holliday, 1972). They feed on, and to a large extent spatially coincide with, juvenile anchovies.

The concentrations of fish in the ocean are generally described in terms of spatial number densities, n_A , which may be derived by integrating volumetric number densities (number/m³) over depth. Average spatial number densities, n_A , of adult juveniles in the south-central region of the Yellow Sea and north-central region of the East China Sea in May and June are typically about 0.7/m² (top view) (Ohshimo, 1996). This value is comparable to the spatial density of sardines in the Gulf of Lion (Diachok 1999) and in the Bristol Channel (Ching and Weston, 1971), where large bioacoustic effects on transmission loss have been reported. According to Fig. 1, values of this parameter in the near-shore spawning regions are much larger than average values.

III. BIOACOUSTICS

Anchovies belong to a class of swim-bladder-bearing fish called physostomes, which have “open” swim bladders (Whitehead and Blaxter, 1989). Such fish maintain air in their swim bladders by swallowing air at the surface. The volumes of their swim bladders decrease with increasing depth, in accord with Boyle’s law. The resonance frequency, f_0 , of physostomes may be derived from an equation originally formulated to describe the resonance frequency a spherical bubble (Minnaert, 1933):

$$f_0 = (1/2\pi r)(3\gamma P/\rho)^{1/2} \quad (1)$$

where r is the radius, P is the hydrostatic pressure, ρ is the density of the surrounding fluid, and γ is the ratio of specific heats of the gas in the bubble. The presence of γ in Eq. (1) implies that the gas behaves adiabatically. This assumption is valid for relatively large bubbles (Clay and Medwin, 1977), and is therefore probably valid for fish swim bladders.

The radius of bubbles in water satisfies Boyle’s law:

$$Pr^3 = P_0r_0^3, \quad (2)$$

where P_0 and r_0 are the pressure and radius at the surface, and

$$P = P_0(1 + 0.1d/d_0), \quad (3)$$

where d is the depth in m and d_0 equals 1 m. Neglect of changes in r due to changes in temperature results in errors, which are smaller than 1%.

The resonance frequency of nonspherical bubbles is higher than the resonance frequency of spherical bubbles, which have the same volume. To account for this effect, Eq. (1) needs to be increased by a factor, ε , which is a function of the eccentricity, e , of the bubble, defined as the ratio of the major and minor axes of an ellipsoid that provides the “best” fit to the shape of the swim bladder. This correction was originally derived for prolate spheroids by Weston (1967). Subsequently Feuillade and Werby (1994) derived corrections for swim bladders modeled as cylinders with hemispherical end caps and prolate spheroids. The difference between calculations of ε based on these models is small (about 3%), which suggests that the correction is minimally affected by the shape of the swim bladder.

Substituting Eqs. (2) and (3) into Eq. (1) and inserting the correction for the eccentricity leads to

$$f_0 = \varepsilon[(1 + d/d_0)^{5/6}/(2\pi r_0)](3\gamma P_0/\rho)^{1/2}, \quad (4)$$

which can be simplified to

$$f_0 = 322\varepsilon(1 + 0.1d)^{5/6}/r_0, \quad (5)$$

where r_0 is the effective radius in cm at the surface and d is the depth in meters. The effective radius is defined as the radius of a sphere of the same volume as the swim bladder. For California anchovies e is approximately equal to 6 (Whitehead and Blaxter, 1989), which corresponds to $\varepsilon=1.14$ for prolate spheroids and 1.18 for a cylinder with hemi-spherical end caps (Feuillade and Werby, 1994). Calculations, based on this equation with $\varepsilon=1.16$, are consistent within $\pm 7\%$ of laboratory (Baltzer and Pickwell, 1970) and field measurements (Holliday, 1972) of f_0 of anchovies (Diachok, 2001). The good agreement between theory and measurements supports the “adiabatic” assumption and the applicability of the theoretical correction for eccentricity.

This equation is valid when the separation between fish, s , is comparable to or larger than the wavelength, λ_0 , at f_0 . When fish are in school and $s \ll \lambda_0$, the resonance frequency of a school is a function of s and the number of fish in the school, N , and is generally less than f_0 . See Feuillade *et al.* (1996) and Diachok (1999) for further discussion of this phenomenon.

Several theories (Andreeva, 1964; Love, 1978; McCartney and Stubbs, 1970; Feuillade and Nero, 1998) have been proposed to account for the effects of the physical properties of the swim bladder membrane on f_0 . However, the good agreement between measurements and calculations of f_0 of anchovies, based on Eq. (5), indicate that the hypothesized effects for this species are small.

The average “standard” length, SL, of anchovies in the vicinity of Qiu *et al.*’s measurements was approximately 10 cm (Ohshimo, 1996) (for anchovies SL equals 0.84 TL, where TL is the total length from nose to tail). Holliday (1978) derived empirical relationships between r_0 and SL of California anchovies from measurements on a large number of samples, which were taken in June and March. Results are summarized in Fig. 2. Measurements made on postlarval juveniles are also shown for comparison. The small difference between r_0 in March and June is probably controlled by the seasonal variation of the “lipid” (oil) content of fish flesh

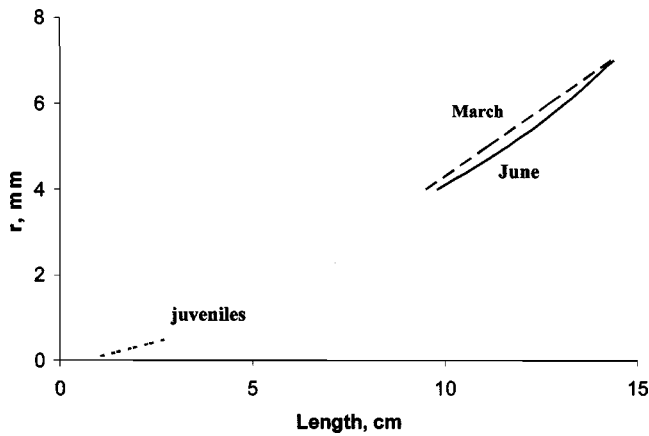


FIG. 2. Effective radius, r_0 , versus standard length, L , of postlarval juvenile, and adult California anchovies (*Engraulis mordax*) in June and March (Holliday, 1978).

(Blaxter and Batty, 1990). Based on Holliday's measurements in June, which will be assumed to be representative of Qiu *et al.*'s measurements in August, the magnitudes of r_0 and e of Japanese anchovies are projected to be approximately 0.40 cm and 6, respectively. The corresponding value of ε is 1.16 ± 0.02 . Assuming these values in Eq. (5) yields f_0 equal to 1.38 kHz at 6 m.

The absorption coefficient due a large number of identical, dispersed fish with swim bladders may be calculated from the following equation (Weston, 1967),

$$\alpha \text{ (dB/m)} = 4.34n\sigma_E = n\lambda_0 r/Q[(1 - f_0^2/f^2)^2 + 1/Q^2], \quad (6)$$

where σ_E is the extinction cross section of a single bubble (due to scattering) in m^2 , n is the number density (number per m^3), $Q = f_0/\Delta f$ (where Δf is the full width at -3 dB relative to the peak of the absorption line), and λ_0 is the wavelength at f_0 . This equation has been employed successfully to derive the number densities of sardines in the Gulf of Lion (Diachok *et al.*, 2001) and the Bristol Channel (Ching and Weston, 1971).

IV. GEO-ACOUSTICS

Qiu *et al.* (1999) conducted transmission loss measurements at the site shown in Fig. 1. The depth was 40 m, the bottom boundary was flat, and the bottom consisted of a thick layer of unconsolidated sediments. Qin and Zhao's (1987) measurements of surface sediments in the Yellow Sea indicate that this site consist of "silty sand" ($\approx 58\%$ sand). Geo-acoustic parameters of this site are not known. Geo-acoustic parameters have been estimated at two nearby, unspecified sites by Zhou (1985), and at the site in the central Yellow Sea shown in Fig. 1 by Dahl *et al.* (1998) and Rogers *et al.* (2000).

Zhou inverted geo-acoustic parameters from propagation experiments at two unspecified sites (A and B) near the coast of China in the Yellow Sea. The sediments at Zhou's site A, where the water depth was 37 m, consisted of 39% sand. Sediments at site B, where the water depth was 29 m, consisted of 29% sand ("sandy silt"). Zhou concluded that the bottom at both sites may be characterized by a half-space

with c_p equal to 1584 m/s, independent of frequency between 100 and 500 Hz. Subsequently, Zhou *et al.* (1987) offered a revised geo-acoustic model that included a low sound speed (c_p equals 1560 m/s at the top), high gradient (5 m/s/m), 10-m-thick, transition layer overlying a high sound speed (1610 m/s) half-space.

Rogers *et al.* and Dahl *et al.*'s site consists of "silty sand." Rogers *et al.* concluded that the bottom at this site may be characterized as a half-space with c_p equal to 1587 m/s, independent of frequency between 100 Hz and 1.5 kHz. Dahl *et al.*, who recorded data at the same site, but on a different hydrophone array, reached a different conclusion, viz., that the bottom may be characterized by a low sound speed (1555 m/s at the top), 2-m-thick, transition layer overlying a high sound speed (1650–1700 m/s) half-space, independent of frequency between 100 and 800 Hz. Both models resulted in good matches between calculations and measurements of transmission loss.

In this paper it will be assumed that the bottom may be characterized by an interfacial sound speed, c_p , and a subsurface gradient. The measurements discussed above suggest the following upper and lower bounds on c_p for inversion calculations: 1700 and 1555 m/s. Hamilton's (1987) review suggests the following upper and lower bounds on α_p (dB/ λ): 1 and 0.02 dB/ λ . Previously reported measurements of α_p in the Yellow Sea and elsewhere will be summarized and compared with inverted values of α_p at Qiu *et al.*'s site in Sec. VIII.

V. TRANSMISSION LOSS MODEL

It will be assumed that sound propagation at the frequencies and ranges of interest may be calculated with a deterministic normal mode model, and that stochastic mode conversion due to internal waves may be ignored. This assumption seems reasonable at the frequencies and ranges of interest (Dozier and Tappert, 1978). The transmission loss model, which will be employed to model the effects of bio-acoustic absorption layers on transmission loss, is based on the coupled normal mode model, C-SNAP (Ferla *et al.*, 1993). This model can account for range-dependent sound speed profiles in the ocean, depth-dependent c_p and depth-independent α_p in the bottom, and scattering loss at boundaries. Ferla's initial modifications to C-SNAP (Diachok and Ferla, 1996) were designed to permit calculation of the effects of two bio-acoustic absorption layers. The resultant code was successfully used to model the effects of bio-acoustic absorption layers on transmission loss in the Gulf of Lion (Diachok, 1999). Subsequently, this model was refined to permit consideration of up to 100 layers. The more recent version of this code, designated BIO-C-SNAP, permits modeling the depth dependence of α_B , including Gaussian and other distributions of α_B .

The absorbing layer will be characterized by a horizontally uniform bio-absorption coefficient, α_B , a mid-layer depth, d , and a layer thickness, t . It will be assumed that the presence of fish in this layer has a negligible effect on sound speed. This is a reasonable assumption when the void fraction, β , which is defined as

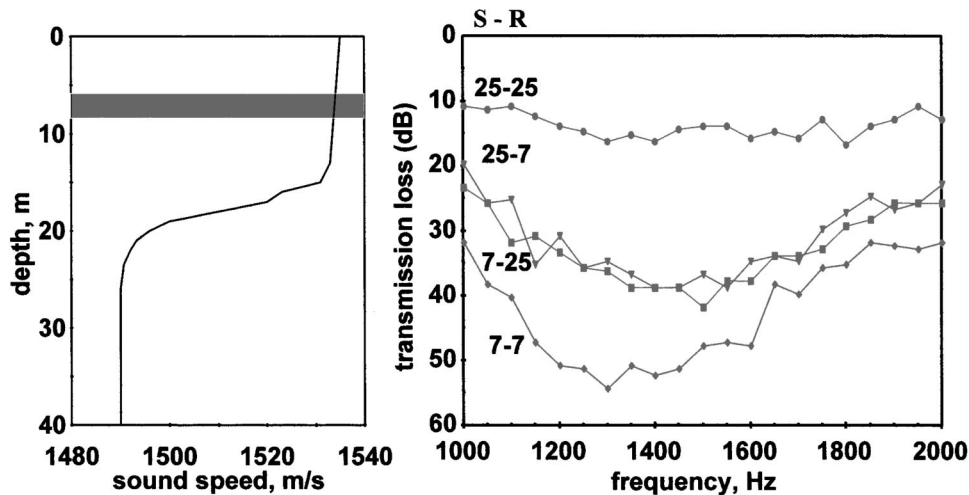


FIG. 3. The postulated depth of the layer of dispersed anchovies (gray) and Qiu *et al.*'s (1999) measurements of the sound speed profile (left) and transmission loss at 5.7 km versus source and receiver depths (right).

$$\beta = (4/3)\pi r^3/s^3, \quad (7)$$

is less than approximately 10^{-6} (Commander and Prosperetti, 1989). Limited photographic data indicate that average separation between anchovies, which are dispersed near the surface at night, are about 10 fish lengths (Aoki and Inagaki, 1988). This is comparable to, but somewhat larger than, the average value of s , which will be derived through comparison of propagation model calculations with Qiu *et al.*'s data, viz., 5 L (which will be discussed in Sec. IX). Assumption of r and s equal 0.34 cm and 50 cm respectively results in β equal to 10^{-6} .

The assumption that the absorption layer is horizontally uniform is reasonable for dispersed anchovies at night. A small percentage of anchovies, however, can remain in schools at night. The presence of schools would violate the assumption of horizontal uniformity. Schools of anchovies consist of highly spatially localized "bubble clouds," which are nominally 3 m thick, 10 m wide, separated by distances on the order of 100 m, and resonate at frequencies which are less than f_0 (Diachok, 1999). The magnitude of β of some schools may be sufficiently low to affect the speed of sound. Inversion calculations, which will be presented in Sec. VII, may be affected by these phenomena, particularly at frequencies below f_0 .

The magnitude of the attenuation coefficient of compressional waves in sediments, α_P , will be assumed to be independent of depth (a limitation of the C-SNAP model). This is not a significant limitation when c_P is greater than the speed of sound in water and the shear speed, c_S , is much less than c_P . c_S will be assumed to equal zero.

VI. TRANSMISSION LOSS MEASUREMENTS

Qiu *et al.* (1999) reported transmission loss measurements at frequencies between 1 and 2 kHz in August at the site shown in Fig. 1. Their experiment was conducted at night, when anchovies are generally dispersed near the surface. Transmission loss measurements (in relative dB) were made between pairs of sources and receivers at 7 and 25 m, which were above and below the base of the mixed layer (about 15 m), as illustrated in Fig. 3. The water depth was 40 m. Pairs of explosives were deployed at these depths at

ranges between 0.4 and 6.6 km (measurements at 6.6 km were made at the beginning of twilight and were excluded from the analysis presented here). Results at a range of 5.7 km are shown in Fig. 3. This figure illustrates the main result of their experiment: (1) relative losses are greatest (~ 40 dB) when the both source and receiver are at a depth of 7 m, compared to losses when source and receiver were at 25 m, and (2) the loss is highly frequency selective, being largest at a frequency of about 1.35 kHz.

Figure 4 provides a comparison of the calculated and measured frequency dependence of the bio-acoustic component of the attenuation coefficient of the transmission channel, A_B , in dB/km. A_B is defined as the excess attenuation above cylindrical spreading. The purpose of this figure is to show that these data are consistent with theoretical calculations of the frequency dependence of bio-absorptivity due to 10 cm long anchovies. These calculations were derived from TL measurements made between source and receiver depths of 7 m. Measured values were derived from the measured rate of change of signal level with range, after correcting for cylindrical spreading, and a small frequency-independent attenuation, A_X , which is hypothetically caused by a combination of α_P and α_B . The latter is hypothetically due to schools of anchovies or a dispersed larger species which resonate at a lower frequency. The theoretical calculations, shown in Fig. 4, were based on Eq. (5). The best fit to the

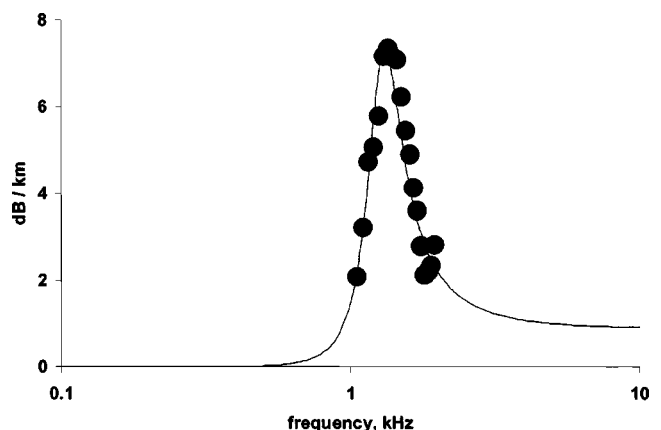


FIG. 4. Measured and calculated spectra of bio-acoustic attenuation.

data was determined by searching through the following parameters: A_B at 1.35 kHz, the Q of the absorption line, and A_X . The results of fitting Qiu *et al.*'s data to this model are consistent with A_B equal to 7.5 dB/km at 1.35 kHz, Q equal to 2.7, and A_X equal to 1 dB/km. The latter is small compared to A_B at 1.35 kHz, which supports the hypothesis that bio-absorptivity due to 10-cm-long anchovies is the dominant cause of attenuation at this frequency. The absorption coefficient of the layer of anchovies, α_B , may be estimated from the calculated value of A_B by assuming that the acoustic field is uniformly distributed in the water column and the following equation (Diachok, 1999):

$$\alpha_B = A_B D / t, \quad (8)$$

where α_B is in dB/ λ , A_B is in dB/ λ , D is the ocean depth in m, and t is the thickness of the layer in m. For $A_B = 7.5 \text{ dB/km} = 0.0075 \text{ dB/m} = 0.0083 \text{ dB}/\lambda$, $D = 40 \text{ m}$, $\lambda_0 = 1.1 \text{ m}$, and $t = 0.3 \text{ m}$ (based on inversion calculations, which will be presented in Sec. VII), $\alpha_B = 1.1 \text{ dB}/\lambda$. This is approximately equal to the value of α_B , which will be derived through comparison of propagation model calculations with data, which will be presented in Sec. VII.

The magnitude of Q is determined by the values of Q_0 , the inherent Q of individual fish, Q_L , the Q associated with the distribution of lengths, and Q_Z , the Q associated with the distribution of depths through a generalized form of Weston's (1967) equation:

$$1/Q^2 = 1/Q_0^2 + 1/Q_L^2 + 1/Q_Z^2. \quad (9)$$

This equation assumes that each of these processes results in a "random" distribution of resonance frequencies. Laboratory measurements indicate that Q_0 of anchovies at 6 m equals approximately 4.5 (Baltzer and Pickwell, 1970). The magnitude of Q_L of adult anchovies in the Yellow Sea is approximately 3.5, based on measurement of the width of the 3 dB down points in Deng and Zhao's (1991) measurements of the length distributions of anchovies in the Yellow Sea, Holliday's empirical relationships between L and r_0 , and Eq. (5). Disregarding the effect of Q_Z , the calculated value of Q is 2.8. This is in excellent agreement with the measured value, 2.7, which suggests that Q_Z is large, i.e., greater than about 7. The magnitude of Q_Z , which may be estimated from Eq. (5), is approximately equal to d/t , where d is the average depth and t is the thickness of the layer. It will be shown in the next section that d at Qiu *et al.*'s site is approximately equal to 6.9 m, which implies that t is less than approximately 1 m.

The wind speed during this experiment was about 10 knots (Qiu, personal communication). Attenuation due to surface waves and subsurface bubbles at frequencies below 2 kHz at 10 knots for a water depth of 40 m is less than 0.1 dB/km (Weston and Ching, 1989). Chemical absorptivity at these frequencies and ranges is also small (Thorpe, 1965). Attenuation due to the rms roughness ($\sim 0.3 \text{ m}$) at the bottom boundary is also small at these frequencies (Zhou, 1985).

According to Holliday's empirical relationships between L and r_0 , and Eq. (5), 10-cm-long anchovies with r_0 equal to 0.4 cm and ε equal to 1.15 resonate at 1.35 kHz at a depth of 5.8 m. The uncertainty in this estimate of depth, $\pm 1 \text{ m}$, was

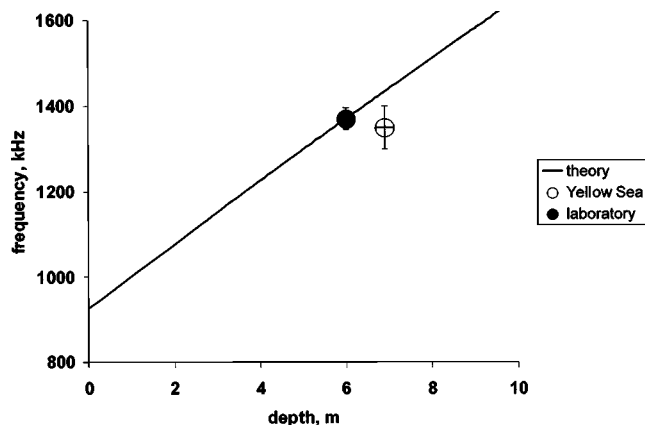


FIG. 5. Measured resonance frequencies due to anchovies in the Yellow Sea (Qiu *et al.*, 1999), and extrapolations from Baltzer and Pickwell's (1970) laboratory measurements.

derived from the estimated uncertainties in the measured value of the average fish length ($\pm 10\%$), the calculated value of r_0 ($\pm 5\%$), and the measured value of f_0 ($\pm 5\%$). These observations were then extrapolated to the Yellow Sea, assuming that the differences in the relationships between L and r_0 and ε of the swim bladders of Japanese and California anchovies are small. This assumption is reasonable, since the magnitude of the correction for eccentricity is small. Figure 5 provides a comparison between calculations and measurements of f_0 of 10-cm-long anchovies in the laboratory and in the Yellow Sea. The laboratory estimate was extrapolated from measurements of f_0 (1.275 kHz), which were made on 10.6-cm-long California anchovies ($r_0 = 0.43 \text{ cm}$) at 6 m. The depth of anchovies in the Yellow Sea was inverted from TL data, as discussed in Sec. VII.

VII. INVERSION OF BIO AND GEO-ACOUSTIC PARAMETERS

Additional evidence in support of the bio-acoustic paradigm will be provided by demonstrating that the depth of the absorption layer, which was derived from Eq. (5), viz., $5.8 \pm 1 \text{ m}$, is consistent with the average layer depth, which will be "inverted" from transmission loss measurements. The inversion method, which will be employed here, is similar to previously reported inversion methods for estimation of ocean and geo-acoustic parameters from sound propagation measurements (Diachok *et al.*, 1995).

TL calculations were based on incoherent addition of modes generated with the BIO-C-SNAP model. Incoherent addition of modes was assumed to result in a reasonable approximation to (0.1 kHz) frequency averaging of Qiu *et al.*'s data (Jensen and Kuperman, 1983).

It was assumed that bio-acoustic absorptivity was due to a single layer, which was modeled with a single uniform (depth independent) absorption coefficient. Incorporation of a "Gaussian" shape did not have a significant effect on calculations. It was assumed that the bottom could be characterized with an interfacial sound speed, c_p , and a subbottom gradient equal to 1.3 m/s/m (Hamilton, 1987). Assumption of other values of this parameter (up to 2 m/s/m) had no effect on inverted parameters. The rms roughness was assumed to

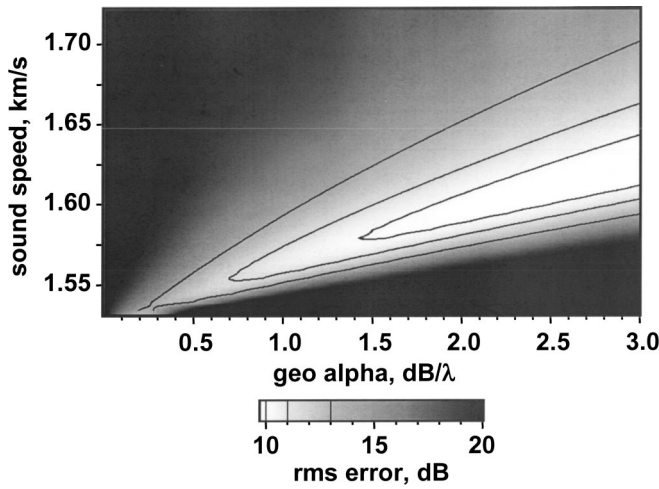


FIG. 6. Ambiguity surface of the rms difference, Δ (in dB), between measured and calculated transmission loss at 1.35 kHz versus bottom sound speed and attenuation coefficient, assuming that all excess attenuation is due to the bottom. The inverted sound speed is 1625 m/s and attenuation coefficient is equal to 9.5 dB/λ. Contours: ± 10 , 11, and 13 dB.

equal zero. Assumption of a small rms roughness (0.3 m) at the bottom interface did not reduce the rms difference between calculations and measurements.

Initial inversion calculations were conducted for the case of no bio-acoustic absorption layer. The search was based on minimizing the rms difference, Δ , between calculated and measured transmission losses as a function of three parameters: c_p , α_p , and the correction for source level. All of Qiu *et al.*'s TL data, which included measurement at ranges between 0.4 and 5.7 km (but excluded data at 6.6 km) and source and receiver depths of 7 and 25 m, were employed in the inversion. The search method was "comprehensive", i.e., Δ was calculated for all combinations of parameters. The resultant ambiguity surface, Fig. 6, shows that the minimum value of Δ was 9.5 dB for c_p equal to 1625 m/s and α_p equal to 3 dB/λ. The unacceptably high value of Δ indicates that the environmental model, which was assumed to generate replica fields, does not provide a good approximation to reality. It is noteworthy that the inverted value of α_p (3 dB/λ) is much larger than Hamilton's (1987) upper bound on this parameter (1 dB/λ) at high frequencies.

The correction for source level had to be included in the search, since Qiu *et al.* did not report absolute values of TL. The search was not sensitive to the magnitude of this unknown: addition of 50 dB to Qiu *et al.*'s published values of relative TL, which resulted in reasonable absolute values of TL at 1 kHz at 0.4 km, did not significantly affect calculated ambiguity surfaces.

Subsequent calculations included a search through both bio and geo parameters. This search was based on minimizing the rms difference, Δ , between calculated and measured transmission losses as a function of six parameters: α_B , d , t , c_p , α_p , and the correction for source level. The search was conducted by varying two to three parameters, while holding others constant. This procedure was repeated several times for all possible combinations of pairs of parameters until Δ approached an asymptotic value. A sufficiently large number of ambiguity surfaces were calculated to ensure that local

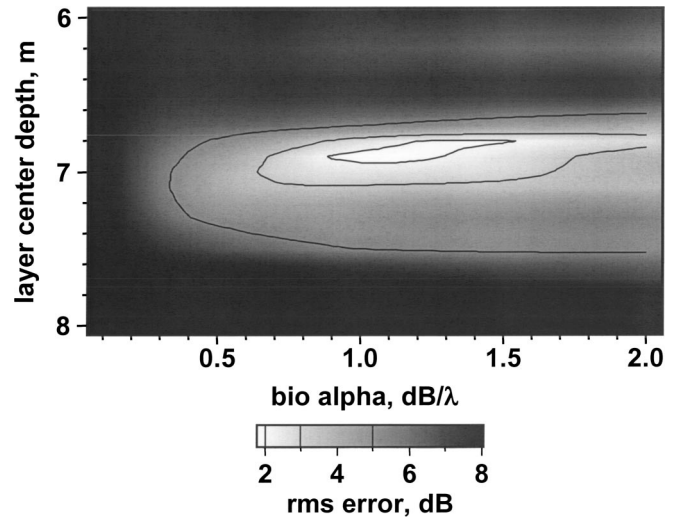


FIG. 7. Ambiguity surface of the rms difference, Δ (in dB), between measured and calculated transmission loss at 1.35 kHz versus layer depth and absorption coefficient within the layer. Layer thickness equals 0.3 m. The inverted layer depth is approximately equal to 6.8 m. Contours: ± 2 , 3, and 5 dB.

minima were avoided and that ambiguity surfaces for all possible combinations of parameters were self-consistent. Figures 7 and 8 show ambiguity surfaces for assumed values of the layer depth, d , layer thickness, t , and the bio-attenuation coefficient, α_B .

Figure 7 shows the magnitude of Δ vs. d and α_B at 1.35 kHz for t equal to 0.3 m and α_p equal to 0.15 dB/λ. The minimum Δ , 1.7 dB, occurs when d equals 6.9 m and α_B equals 1.1 dB/λ. This value of α_B is virtually identical to the heuristic estimate presented in Sec. VI. This figure indicates that Δ is less than 3 dB (within 1.3 dB of the minimum) when d is between 6.6 and 7.0 m, and α_B is between 0.9 and 1.4. The inverted value of d is consistent with Eq. (5) based estimate (5.8 ± 1 m).

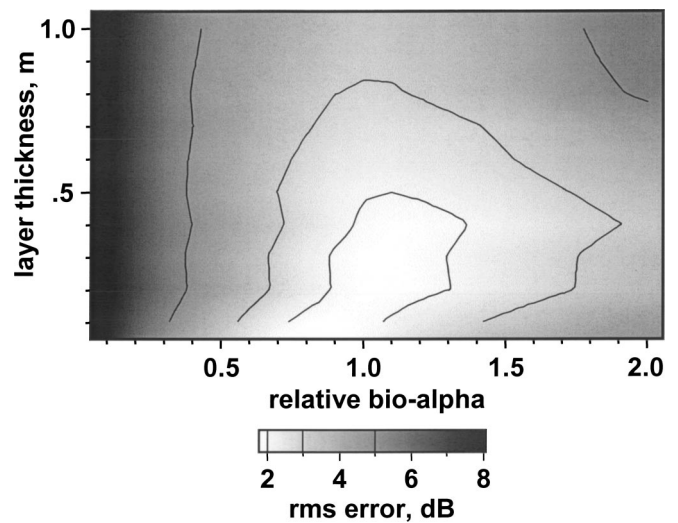


FIG. 8. Ambiguity surface of the rms difference, Δ (in dB), between measured and calculated transmission loss at 1.35 kHz versus layer thickness, t , and "relative bio-alpha" within the layer, α_b . Bio-alpha, α_B , may be calculated from the equation, $\alpha_B = 0.3 \times \alpha_b / t$. Layer depth is 6.8 m. The inverted layer thickness is approximately equal to 0.3 m. Contours: ± 2 , 3, and 5 dB.

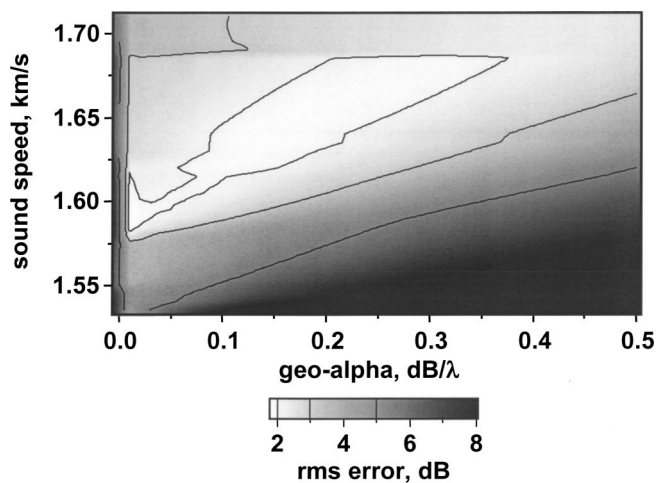


FIG. 9. Ambiguity surface of the rms difference, Δ (in dB), between measured and calculated transmission loss at 1.35 kHz versus bottom sound speed and attenuation coefficient. The inverted sound speed is 1645 m/s and attenuation coefficient is equal to 0.19 dB/λ. Contours: ± 2 , 3, and 5 dB.

Figure 8 shows the magnitude of Δ vs. t and α_B for a layer depth of 6.9 m and α_P equal to 0.19 dB/λ. In this case the minimum Δ , 1.8 dB, occurs when t equals 0.3 m. This figure indicates that Δ is less than 3 dB (within 1.2 dB of the minimum) when t is between 0.0 and 0.8 m. The relatively small inverted value of t is consistent with previously reported measurements of the thickness of plankton layers in the ocean. The thickness of fish layers is hypothetically driven by the thickness of plankton layers. The latter are frequently less than 1 m thick (Cowles *et al.*, 1998; Holliday, 1995). These observations are consistent with limited, previously reported echo sounder measurements of t of anchovies (Barange *et al.*, 1996; Holliday, personal communication), myctophids (Benoit-Bird and Au, 2004) and juvenile pollock (Francis and Bailey, 1983). Barange *et al.*'s measurements, which were made over several months and over hundreds of km, resulted in a time-averaged value of t of 2 m. This suggests that measurements, which were made over shorter durations, such as one night, and shorter distances, such as a few km, were less than 2 m. Benoit-Bird and Au reported that t varies with the phase of the moon and is minimum, 1 m, at night during a full moon. Holliday's unpublished measurements (personal communication) suggest that some fish (probably anchovies) aggregate at thin plankton players. Francis and Bailey (1983) reported that the layer thickness of juvenile pollock was less than 2 m when there was a strong thermocline.

Figure 9 shows the magnitude of Δ vs. c_P and α_P for fixed values of bio-acoustic parameters. This figure illustrates the low sensitivity of the data at 1.35 kHz to geo-acoustic parameters. The magnitude of Δ was less than 3 dB (within 1.3 dB of the minimum for c_P between 1580 and >1720 m/s, and α_P between 0.02 and >0.5 dB/λ; the magnitude of Δ was less than 2 dB (within 0.3 dB of the minimum) for c_P between 1580 and 1680 m/s, and α_P between 0.02 and approximately 0.36 dB/λ. The best fit occurred at c_P equal to approximately 1645 m/s and α_P equal to approximately 0.19 dB/λ, where Δ equaled 1.7 dB.

Figure 10 shows calculated magnitudes of Δ vs. α_P and

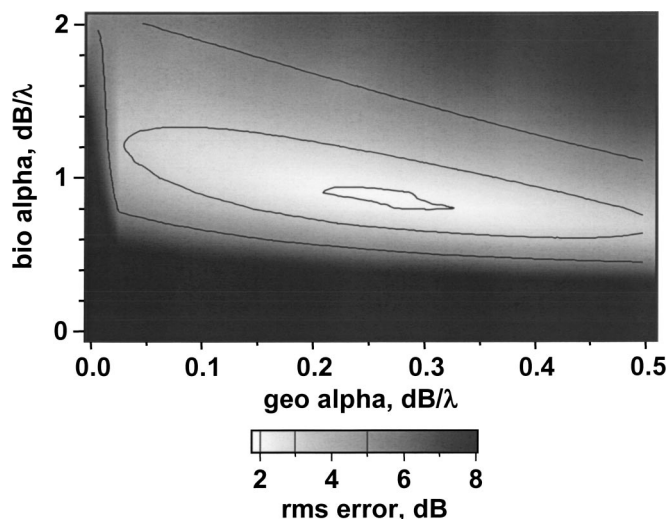


FIG. 10. Ambiguity surface of the rms difference, Δ (in dB), between measured and calculated transmission loss at 1.35 kHz versus geo alpha and bio alpha. The inverted values of α_P and α_B are equal to 0.3 and 1 dB/λ, respectively. Contours: ± 2 , 3, and 5 dB.

α_B for fixed values of c_P (1645 m/s), t (0.3 m), and d (6.9 m) at 1.35 kHz. The magnitude of Δ was less than 3 dB for α_B between 0.6 and 1.3 dB/λ, and α_P between 0.03 and 0.5 dB/λ. The magnitude of Δ was less than 2 dB for α_B between 0.8 and 0.9 dB/λ, and α_P between 0.22 and 0.32 dB/λ. The optimum values of α_P and α_B equal 0.27 and 0.9 dB/λ, consistent with results shown in Figs. 7–9.

Calculations at 2 kHz resulted in an acceptable value of Δ , 2 dB. Inversion of bio-acoustic parameters, which assumed the same values of geo-acoustic parameters that were derived at 1.35 kHz, resulted in values of d , t , and α_B , which were consistent with values derived at 1.35 kHz. The inverted value of α_B was scaled down by a factor of about 0.3, in accord with Eq. (5). Inversion of c_P , however, was not successful at this frequency. The optimum value of c_P was at the upper bound of the search space, 1720 m/s, where Δ was 2 dB, and the 3-dB bounds on c_P exceeded the search space. The low sensitivity to c_P at this frequency may have been caused by shortcomings of the geo-acoustic model, or by stochastic mode coupling due to internal waves. The latter increases with increasing range and frequency (Dozier and Tappert, 1978; Duda and Preisig, 1999). Calculations at 1 kHz resulted in an unacceptably high value of Δ , 3.5 dB, hypothetically due to shortcomings of the geo- or bio-acoustic models.

A comparison of the minimum values of Δ for inversions that assumed only geo parameters, and a combination of bio and geo-acoustic parameters, is shown in Fig. 11. These results indicate that inclusion of a bio-absorption layer results in much smaller values of Δ at all frequencies, despite the apparent shortcomings of the geo-acoustic model at 1 and 2 kHz.

VIII. GEO-ALPHA

The purpose of this section is to provide a context for comparison of the value of α_P (dB/λ) inverted from Qiu *et al.*'s data with previously reported measurements of this

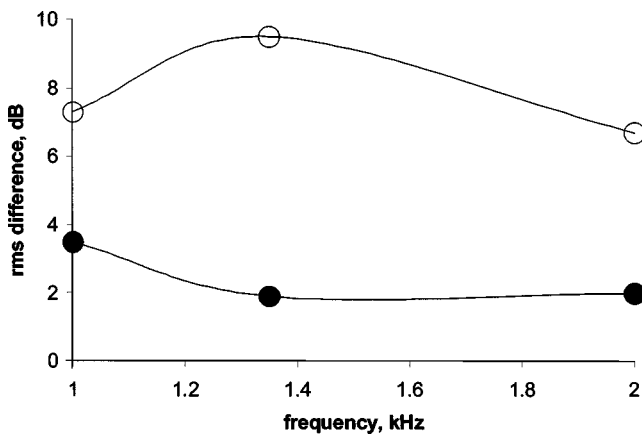


FIG. 11. Minimum rms difference, Δ (in dB), between measured and calculated transmission loss, assuming attenuation is due to only geo parameters (○), and bio and geo parameters (●).

parameter. Figure 12 shows the inverted values of α_p at Qiu *et al.*'s site, Zhou's estimates at a nearby, unspecified site, Dahl *et al.* and Rogers *et al.*'s estimates at the site in the central Yellow Sea depicted in Fig. 1, and analytical calculations based on work reported by LeBlanc *et al.* (1992), Turgut *et al.* (2002), and Buckingham (2000).

LeBlanc *et al.* (1992) fit "chirp" sonar measurements at frequencies between about 3 and 10 kHz with an analytical model, which assumed that α_p was due to a relaxation mechanism. According to this model, α_p increases with frequency to the first power at frequencies below the relaxation frequency. Turgut *et al.* fit Hamilton's (1972) *in situ* measurements at frequencies between 3 and 25 kHz with theoretical calculations based on the "Biot" model. According to this model, α_p also increases with frequency to the first power at low frequencies. Buckingham's (2000) theory assumes that α_p is controlled by intergranular friction. According to his model, α_p is independent of frequency. The lines drawn in Fig. 12 assume that both mechanisms contribute, and are described by the equation

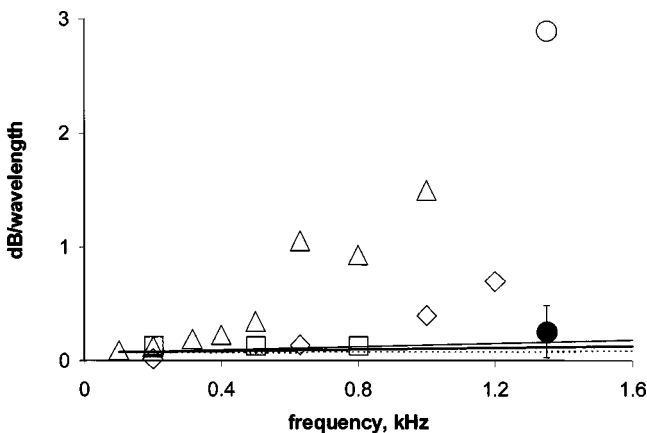


FIG. 12. Values of α_p , which were inverted from Qiu *et al.*'s data at 1.35 kHz, based on replica fields which incorporate bio- and geo-alpha (●) and only geo-alpha (○), and Zhou's (1987) (Δ), Rogers *et al.* (2000) (◇), and Dahl *et al.*'s (1998) (□) measurements of α_p in the Yellow Sea, which assume that all attenuation is due to geo-alpha, and analytical calculations in "sand" (—) and "silt" (---) from LeBlanc *et al.* (1992), and "fine sand" (—) and "clayey silt" (-.-) from Turgut *et al.* (2002).

$$\alpha_p = bf + h, \quad (10)$$

where α_p is in dB/λ, f is the frequency in kHz, h equals 0.07, and b equals 0.011 for "silt" and 0.031 for "sand" from LeBlanc *et al.*'s calculations, and 0.012 for "clayey silt" and 0.072 for "fine sand" from Turgut *et al.*'s calculations. The first term accounts for the relaxation mechanism at low frequencies. The second term is consistent with Buckingham's theory; the magnitude of h was selected to fit Dahl *et al.*'s, Rogers *et al.*'s, and Zhou's data at 200 Hz.

The inverted value of α_p at 1.35 kHz at Qiu *et al.*'s site, based on the bio and geo model, is consistent with both Turgut *et al.*'s and LeBlanc *et al.*'s extrapolations from high-frequency measurements. This is a reasonable result, since the magnitude of α_p of "silty sand" lies between "sand" and "silt" at high frequencies (Hamilton, 1972).

Values of α_p inverted from propagation measurements are sensitive to the assumed sound speed profile in the bottom. Zhou's estimates of α_p , which assumed that the bottom may be characterized by a half-space (1584 m/s), increase with frequency as shown in Fig. 12. Zhou *et al.*'s (1987) revised estimates of α_p (not shown) from Zhou's data, which assumed a low sound speed, 10-m-thick, transition layer overlying a high sound speed half-space, were about 0.6 times lower than Zhou's original estimates. Zhou *et al.* (1987) also concluded that their previously reported measurements were consistent with α_p equal to $0.45e^{-1.3z}$ dB/λ, where z is in m in the top 10 m. Dahl *et al.* inverted the following values of α_p : 0.2 dB/λ in a low sound speed, 2-m-thick, transition layer, and 0.07 dB/λ in the underlying high sound speed half-space (the average value, 0.14 dB/λ, is shown in Fig. 12). Rogers *et al.*'s estimates of α_p , which assumed that the bottom may be modeled by a half space, increased with frequency, as shown in Fig. 12.

Zhou's estimates and Dahl *et al.*'s and Rogers *et al.*'s estimates of α_p diverge at frequencies above about 500 Hz. Zhou (1985) does not offer an explanation for the abrupt increase in the apparent value of α_p at 630 Hz at his site. This anomaly may be due to α_B . Unfortunately it is not possible to investigate this hypothesis by inverting the depth of the absorbing layer from Qiu *et al.*'s data at 630 Hz, since their data is restricted to frequencies between 1 and 2 kHz, or by testing compliance of Zhou's data with Eq. (5), since the frequency resolution in his data is too low. Furthermore, the apparent values of α_p at frequencies above 1 kHz, which were derived by Rogers *et al.*, may have been biased by a finite value of α_B . According to Ohshimo (1996), the concentration of anchovies diminishes with distance from spawning grounds, and is smaller at the latter site than at Qiu *et al.*'s site (but is not equal to zero as Fig. 1 suggests).

Table I provides a comparison of inverted geo-acoustic parameters at 1.35 kHz, based on results shown in Figs. 7–10, and previously published measurements discussed above. Both 2- and 3-dB bounds are provided for consideration. The historical estimates of c_p are based on Dahl *et al.*'s and Rogers *et al.*'s measurements in "silty sand" (the bottom at Qiu *et al.*'s site consists of "silty sand"). The historical estimate of α_p is based on the analytical extrapolation between Dahl *et al.*'s and Rogers *et al.*'s measurements at

TABLE I. Comparison of inverted and other estimates of bio- and geo-acoustic parameters at 1.35 kHz. The -2 - and -3 -dB bounds on inverted values of d , t , c_p , and α_p are based on contours shown in Figs. 7–10, respectively.

	-3 dB	-2 dB	“Truth”
d (m)	6.6–7.0	6.7–6.9	5.8 ± 1 [Eq. (5)]
t (m)	0.0–0.8	0.3–0.6	< 1 m [Eq. (9)]
c_p (km/s)	1.58– >1.72	1.58–1.68	$1.56^a, 1.59^b$
α_p (dB/ λ)	0.03–0.5	0.22–0.32	$0.10 \pm 0.05^{a-d}$

^aDahl *et al.* (1998).

^bRogers *et al.* (2000).

^cLeBlanc *et al.* (1992).

^dTurgut *et al.* (2002).

low frequencies and LeBlanc *et al.*'s and Turgut *et al.*'s calculations for “sand” and “silt.” Inverted values of c_p and α_p are consistent with these estimates. Table I also provides a comparison of inverted and calculated bio-acoustic parameters, which will be discussed in the next section.

IX. BIO-ACOUSTIC PARAMETERS

The purpose of this section is to compare inverted values of bio-acoustic parameters with previously reported measurements. The depth of the absorption layer, 6.9 ± 0.3 m, was consistent with theoretical calculations of the depth (5.8 ± 1 m) of 10-cm-long anchovies, the dominant species in the Yellow Sea, and with the extrapolated value of the measured resonance frequencies of anchovies in the laboratory; the inverted layer thickness, 0.3 ± 0.2 m, was consistent with measurements of the Q of this absorption line. It is noteworthy that this value of t is also consistent with the nominal thickness of “thin” layers of phyto and zooplankton on continental shelves, which tend to occur when wind speeds are low (Cowles, 1998). The wind speed during Qiu *et al.*'s experiment was about 10 knots.

The inverted value of α_B is consistent with photographic measurements (Aoki and Inagaki, 1988) of the average separation between adult anchovies at night, viz., s equals 10 ± 2 L. Substitution of $\alpha_B = 0.1$ nepers/m, $r = 0.0034$ m, $\lambda_0 = 1.11$ m, and $Q = 2.7$ into Eq. (6) results in n equal to $9.8/\text{m}^3$. This value of n corresponds to s equal to 47 cm, which equals 4.3 L. This estimate assumes that the layer thickness was 0.3 m. If the layer thickness was actually 0.8 m (the upper 3 dB bound on t , shown in Fig. 8), then the n would have equaled $3.7/\text{m}^3$, and s would have equaled 6 L. The latter is in good agreement with Aoki and Inagaki's data.

The inverted value of α_B is significantly higher than historical measurements of the number density of anchovies in the Yellow Sea. Multiplying n by t (0.3 m) results in n_A (the spatial density) equal to $3/\text{m}^2$. The uncertainty in this estimate, approximately $\pm 1/\text{m}^2$, is driven primarily by the uncertainties in Q and α_B (both are approximately $\pm 20\%$). This estimate can be compared with Ohshimo's (1996) estimate of the *average value* of this parameter ($0.6/\text{m}^2$) in the central region of the southern Yellow Sea and the central region of the northern East China Sea in May and June. Ohshimo's measurements show that n_A increases with proximity to the spawning grounds off China, qualitatively consistent with spatial distributions shown in Fig. 1. Further-

more, n_A in the vicinity of eddy-associated “habitats” can be much larger (by a factor of about 4) than average values. Consequently, the actual number density in the vicinity of Qiu *et al.*'s site may have been significantly greater than the average value in the Yellow Sea. In addition, actual *average* number densities in the Yellow Sea may have been higher than Ohshimo's estimates by a factor of about 2, due to his neglect of the depth dependence of target strength, avoidance, and attenuation due to near surface bubbles (Diachok *et al.*, 2001).

X. UNCERTAINTY OF INVERTED PARAMETERS

We consider the inversion at 1.35 kHz as “valid,” because the inverted bio and geo-acoustic parameters were consistent with previously reported estimates. In contrast, we consider the inversion at 1 kHz “not valid,” because the value of Δ , 3.5 dB, is too high; we also consider the inversion at 2 kHz “not valid” because the inverted value of c_p is much greater than previously reported estimates.

Knobles *et al.* (2003) provide a good description of the sources of uncertainty associated with inversion of geo-acoustic parameters from transmission loss data in shallow water. The uncertainty associated with the validity of the assumed geo-acoustic model is one of the primary determinants of the uncertainty in estimated geo-acoustic parameters. Inversion methods can provide valid estimates of environmental parameters, provided that the environmental model and its associated search space provide a “sufficiently realistic” representation of the environment. The success of this method at 1.35 kHz suggests that the inversion at this frequency was conducted in the context of “sufficiently realistic” bio and geo-acoustic models. In contrast, inversion calculations at higher and lower frequencies were not successful, which suggests that the assumed geo-acoustic model was probably not “sufficiently realistic.” Hypothetically, the lower values of α_B at the higher and lower frequencies resulted in a larger amount of energy in high-order modes, which are more sensitive to the details of subbottom structure (layering, gradients, roughness at interfaces, etc.). This suggests the need for a more complex geo-acoustic model and a larger number of parameters in the search space when α_B is small.

Perhaps a thin transition layer overlying a higher sound speed, lower α_p , as postulated by Zhou *et al.* (1987) and Dahl *et al.* (1998), may have provided a more realistic geo-acoustic model. Or, there may have been two or three or n distinct layers. To preclude model-associated uncertainties, experiments designed to measure effects of bio and geo-alpha on TL should include (1) co-located geo-acoustic measurements, viz., chirp sonar, *in situ* probes and/or cores, to facilitate formulation of a realistic model of subbottom structure, and (2) coincident fisheries echo sounders and direct sampling of fish (trawling, purse seining) to facilitate formulation of a realistic geometrical model of bio-absorption layers.

The uncertainty in source levels of shots is about 2 dB; this places a lower bound on Δ . In addition, temporal and spatial fluctuations in sound speed due to internal waves result in stochastic mode coupling of propagated signals,

which affects the accuracy of predicted replica fields. Furthermore, there is an uncertainty in the depth of the measurement site, and the depths of sources and receivers. An additional small, but possibly significant, source of uncertainty is the accuracy of the propagation model. As a result of these uncertainties, there is no universally accepted formula for calculating the uncertainty of inverted parameters, even when the assumed geo-acoustic model may be considered a valid representation of the environment.

XI. CONCLUSIONS

In summary, inversion calculations at the resonance frequency of anchovies, which assumed that attenuation was due to a combination of bio and geo-alpha, resulted in acceptable values of Δ and realistic values of bio and geo-acoustic parameters, whereas calculations which assumed that all excess attenuation was due to the bottom resulted in unacceptably large values of Δ and unrealistic values of α_p . These calculations assumed relatively simple bio and geo-acoustic models, viz., a biological layer with unknown depth, thickness, and bio-alpha, and a sediment layer with unknown interfacial sound speed, subbottom gradient, and geo-alpha. Inversion at the resonance frequency, where bio-alpha was large, was successful *despite the simplicity of these models*, whereas inversions at higher and lower frequencies, where bio-alpha was small, were not fully successful, hypothetically because of shortcomings of the geo-acoustic model.

The scientifically most important result of this exercise was that we correctly inverted the depth, thickness, and number density of the bio-layer at 1.35 kHz. Inversion of layer depth offers hope of classification of bio-absorption lines based on their depth, an alternative to classification of bio-absorption lines by changes in resonance frequencies at twilight (Diachok, 1999). This is important since the juveniles of most species, and many physoclists, do not undergo significant vertical migration at twilight.

Concurrent inversion of geo-acoustic parameters was a prerequisite for establishing credibility of inverted bio-acoustic parameters. The high sensitivity of this method to uncertainties in geo-acoustic parameters, when bio-alpha is small (or zero), suggests that future bio-absorptivity experiments incorporate independent measurements of geo-acoustic parameters. To minimize effects of internal waves on stochastic mode coupling, experiments should be conducted at shorter ranges.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research Biology and Acoustics Programs and the Naval Research Laboratory (NRL). The authors would like to thank Dr. X. Qiu, formerly of the Shanghai Acoustics Laboratory, Shanghai, China, for providing information about environmental conditions during his experiment and for his permission to reproduce two of his figures. The authors would also like to thank Dr. Tom Hayward and Dr. T. C. Yang of NRL, and Dave Harvey and Allen Lowrie of the Naval Oceanographic Office, for reviewing the manuscript, and Dr. Marshall Orr, Head of the Signal Processing Branch, Dr. Ed

Franchi, Superintendent of the Acoustics Division, and Dr. Eric Hartwig, Associate Technical Director of NRL, for their support and encouragement. O.D. would like to thank Dr. Paul Smith of the Southwest Fisheries Science Center, National Marine Fisheries Service, and Dr. Van Holliday of BAE SYSTEMS Corp. for many helpful discussions. We would also like to thank the formal reviewers, who provided insightful critiques and helpful suggestions.

- Andreeva, I. (1964). "Scattering of sound by air bladders of fish in deep water," *Sov. Phys. Acoust.* **10**, 17–20.
- Aoki, I., and Inagaki, T. (1988). "Photographic observations on the behavior of Japanese anchovy *Engraulis japonica* at night in the sea," *Marine Biol. Prog. Ser.* **43**, 213–221.
- Baltzer, W., and Pickwell, G. (1970). "Resonant acoustic scattering from gas bladder fishes," in *Proceedings of an International Symposium on Biological Scattering in the Ocean*, edited by G. Farquhar (U.S. Government Printing Office, Washington, DC).
- Barange, M., Hampton, I., and Soule, M. (1996). "Empirical determination of *in situ* target strengths of three loosely aggregated pelagic fish species," *ICES J. Mar. Sci.* **53**, 225–232.
- Blaxter, J., and Batty, R. (1990). "Swimbladder behaviour and target strength," *Rapp. P.-V. Reun.-Cons. Int. Explor. Mer* **189**, 233–244.
- Benoit-Bird, K. J., and Au, W. W. L. (2004). "Fine-scale diel migration dynamics of an island-associated sound scattering layer," *Deep-Sea Res., Part I* **51**, 707–719.
- Buckingham, M. J. (2000). "Wave propagation, stress relaxation, and grain to grain shearing in saturated, unconsolidated marine sediments," *J. Acoust. Soc. Am.* **108**, 2796–2815.
- Ching, P., and Weston, D. (1971). "Wide band studies of shallow water acoustic attenuation due to fish," *J. Sound Vib.* **18**, 499–510.
- Clay, C. S., and Medwin, H. (1977). *Acoustical Oceanography* (Wiley, New York).
- Commander, K. W., and Prosperetti, A. (1989). "Linear pressure waves in bubbly liquids: Comparison between theory and experiments," *J. Acoust. Soc. Am.* **85**, 732–746.
- Cowles, T. J., Desiderio, R. A., and Carr, M. E. (1998). "Small scale planktonic structure: persistence and trophic consequences," *Oceanography* **11**, 4–9.
- Dahl, P. H., Eggen, C. J., Tang, D. J., and Spindel, R. C. (1998). "Low frequency sound propagation in the Yellow Sea Results from the 1996 China U.S. experiment," APL UW Technical Report 9804 (Applied Physics Laboratory, Seattle, WA).
- Deng, J., and Zhao, Q. (eds.) (1991). *Marine Fisheries Biology* (in Chinese) (Agriculture Press, Beijing), pp. 357–484.
- Diachok, O. (1999). "Effects of absorptivity due to fish on transmission loss in shallow water," *J. Acoust. Soc. Am.* **105**, 2107–2128.
- Diachok, O. (2001). "Interpretation of the spectra of energy scattered by dispersed anchovies," *J. Acoust. Soc. Am.* **110**, 2917–2923.
- Diachok, O., and Ferla, C. (1996). "Measurement and simulation of the effects of absorptivity due to fish on transmission loss in shallow water," *Proceedings of the Oceans 1996 Conference, MTS/IEEE*, Vol. 2, pp. 524–529.
- Diachok, O., Liorzou, B., and Scalabrin, C. (2001). "Estimation of the number density of fish from resonance absorptivity and echo sounder data," *ICES J. Mar. Sci.* **58**, 137–153.
- Diachok, O., Wales, S., and Smith, P. (2004). "Classification of bio-absorption lines at low frequencies: The Van Holliday connection," *J. Acoust. Soc. Am.* **115**, 8.
- Diachok, O., Caiti, A., Gerstoft, P., and Schmidt, H. (eds.) (1995). *Full Field Inversion Methods in Ocean and Seismo-Acoustics* (Kluwer, Dordrecht, The Netherlands).
- Dozier, L. B., and Tappert, F. D. (1978). "Statistics of normal mode amplitudes in a random ocean. II. Computations," *J. Acoust. Soc. Am.* **64**, 533–547.
- Duda, T. F., and Preisig, J. C. (1999). "A modeling study of acoustic propagation through moving shallow water solitary wave packets," *IEEE J. Ocean. Eng.* **24**, 16–32.
- Ferla, C., Porter, M., and Jensen, F. (1993). Report SM-274, NATO Saclant Undersea Research Centre, La Spezia, Italy.
- Feuillade, C., and Nero, R. W. (1998). "A viscous-elastic swimbladder

- model for describing enhanced-frequency resonance scattering from fish," *J. Acoust. Soc. Am.* **103**, 3245–3255.
- Feuillade, C., and Werby, M. (1994). "Resonances of deformed gas bubbles in liquids," *J. Acoust. Soc. Am.* **96**, 3684–3692.
- Feuillade, C., Love, R., and Nero, R. (1996). "A low frequency acoustic scattering model for small schools of fish," *J. Acoust. Soc. Am.* **99**, 196–208.
- Francis, R. C., and Bailey, K. M. (1983). "Factors affecting recruitment of selected gadoids in the northeast Pacific and East Bearing Sea," in *Year to Year, Interannual Variability of the Environment and Fisheries of the Gulf of Alaska and the Eastern Bering Sea*, edited by W. S. Wooster, Washington Sea Grant Publication, Seattle, WA.
- Hamilton, E. L. (1972). "Compressional wave attenuation in marine sediments," *Geophysics* **37**, 620–646.
- Hamilton, E. L. (1987). "Acoustic properties of sediments," in *Acoustics and the Ocean Bottom*, edited by A. Lara-Saenz, C. Ranz-Guerra, and C. Carbo-Fite (CSIC, Madrid, Spain), pp. 3–58.
- Holliday, D. V. (1972). "Resonance structure in echoes from schooled pelagic fish," *J. Acoust. Soc. Am.* **51**, 1322–1332.
- Holliday, D. V. (1978). *Data analysis and summary, MORDAX II*, Tracor Document No. T-78-SD-002-U, Tracor, Inc. (presently BAE SYSTEMS), San Diego, CA.
- Holliday, D. V. (1995). "Applications of advanced acoustic technology in large marine ecosystem studies," in *Large Marine Ecosystems: Stress, Mitigation, and Sustainability*, edited by K. Sherman, L. Alexander, and B. Gold (AAAS, Washington, DC), Publication 92-39S, pp. 301–319.
- Iverson, S. A., Zhu, D., Johannessen, A., and Toresen, R. (1993). "Stock size, distribution and biology of anchovy in the Yellow Sea and East China Sea," *Fish. Res.* **16**, 147–163.
- Jensen, F. B., and Kuperman, W. A. (1983). "Optimum frequency of propagation in shallow water environments," *J. Acoust. Soc. Am.* **73**, 812–819.
- Knobles, D. P., Koch, R. A., Thompson, L. A., Focke, K. C., and Eisman, P. E. (2003). "Broadband sound propagation in shallow water and geoacoustic inversion," *J. Acoust. Soc. Am.* **113**, 205–222.
- LeBlanc, L. R., Panda, S., and Schock, S. G. (1992). "Sonar attenuation modeling for classification of marine sediments," *J. Acoust. Soc. Am.* **91**, 116–126.
- Love, R. (1978). "Resonant scattering by swimbladder bearing fish," *J. Acoust. Soc. Am.* **64**, 571–580.
- McCartney, B. S., and Stubbs, A. R. (1970). "Measurement of the target strength of fish in the dorsal aspect, including swim bladder resonance," in *Proceedings of an International Symposium on Biological Scattering in the Ocean*, edited by G. Farquhar (U.S. Government Printing Office, Washington, DC).
- Minnaert, F. (1933). "On musical air bubbles and the sounds of running water," *Philos. Mag.* **16**, 235–248.
- Ohshimo, S. (1996). "Acoustic estimation of biomass and school character of anchovy *Engraulis japonicus* in the East China Sea and the Yellow Sea," *Fisheries Sci.* **62**, 344–349.
- Qin, Y. S., and Zhao, S. L. (eds.) (1987). *Late Quaternary Sea-Level Changes* (China Ocean Press, Beijing).
- Qiu, X. *et al.* (1999). "Frequency selective attenuation of sound propagation and reverberation in shallow water," *J. Sound Vib.* **220**, 331–342.
- Rogers, P. H., Zhou, J. X., Zhang, X. Z., and Li, F. (2000). "Seabottom acoustic parameters from inversion of Yellow Sea experimental data," in *Experimental Acoustic Inversion Methods for Exploration of the Shallow Water Environment*, edited by A. Caiti, J. P. Hermand, S. M. Jesus, and M. B. Porter (Kluwer, Dordrecht).
- Thorpe, W. (1965). "Deep ocean sound attenuation in the sub- and low kilocycle per second region," *J. Acoust. Soc. Am.* **38**, 648–664.
- Turgut, A., McCord, M., Newcomb, J., and Fisher, R. (2002). "Chirp sonar sediment characterization at the northern Gulf of Mexico Littoral Acoustic Demonstration Center experimental site," *Proceedings of the Oceans 2002 Conference, MTS/IEEE*, Vol. 4, pp. 2248–2252.
- Weston, D. (1967). "Sound propagation in the presence of bladder fish," in *Underwater Acoustics*, edited by V. Albers (Plenum, New York), Vol. 2, pp. 55–58.
- Weston, D., and Ching, P. (1989). "Wind effects in shallow water acoustic transmission," *J. Acoust. Soc. Am.* **86**, 1530–1545.
- Whitehead, P., and Blaxter, J. (1989). "Swimbladder form in clupeoid fishes," *Zool. J. Linnean Soc.* **97**, 299–372.
- Zhou, J. X. (1985). "Normal mode measurements and remote sensing of sea bottom sound velocity and attenuation in shallow water," *J. Acoust. Soc. Am.* **78**, 1003–1009.
- Zhou, J. X., Zhang, X., Rogers, P. H., and Jarzynski, J. (1987). "Geoacoustic parameters in a stratified sea bottom from shallow water acoustic propagation," *J. Acoust. Soc. Am.* **82**, 2068–2074.

Long range acoustic imaging of the continental shelf environment: The Acoustic Clutter Reconnaissance Experiment 2001

Purnima Ratilal,^{a),b)} Yisan Lai, Deanelle T. Symonds, and Lilimar A. Ruhlmann
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

John R. Preston
Applied Research Lab, Pennsylvania State University, State College, Pennsylvania 16804

Edward K. Scheer
Woods Hole Oceanographic Institute, Woods Hole, Massachusetts 02543

Michael T. Garr
Naval Undersea Warfare Center, Newport, Rhode Island 02841

Charles W. Holland
Saclant Undersea Research Center, San Bartolomeo, La Spezia, Italy

John A. Goff
University of Texas Institute for Geophysics, Austin, Texas 78759

Nicholas C. Makris
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 22 August 2003; revised 6 August 2004; accepted 7 August 2004)

An active sonar system is used to image wide areas of the continental shelf environment by long-range echo sounding at low frequency. The bistatic system, deployed in the STRATAFORM area south of Long Island in April–May of 2001, imaged a large number of prominent clutter events over ranges spanning tens of kilometers in near real time. Roughly 3000 waveforms were transmitted into the water column. Wide-area acoustic images of the ocean environment were generated in near real time for each transmission. Between roughly 10 to more than 100 discrete and localized scatterers were registered for each image. This amounts to a total of at least 30 000 scattering events that could be confused with those from submerged vehicles over the period of the experiment. Bathymetric relief in the STRATAFORM area is extremely benign, with slopes typically less than 0.5° according to high resolution (30 m sampled) bathymetric data. Most of the clutter occurs in regions where the bathymetry is locally level and does not coregister with seafloor features. No statistically significant difference is found in the frequency of occurrence per unit area of repeatable clutter inside versus outside of areas occupied by subsurface river channels. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1799252]

PACS numbers: 43.30.Vh, 43.30.Pc [WLS]

Pages: 1977–1998

I. INTRODUCTION

A long-range bistatic sonar system was used to rapidly image wide areas of the New Jersey continental shelf environment south of Long Island in a field experiment from 27 April to 5 May 2001.¹ The system consisted of a horizontally towed receiving array and two low-frequency vertical source arrays. Source signals were transmitted over long ranges to image scatterers up to tens of kilometers away. Waveguide scattering^{2–6} and propagation⁷ determine the performance of this remote sensing technology.

The field experiment, known as the 2001 Acoustic Clutter Reconnaissance Experiment (ACRE),¹ is a part of the U.S. Office of Naval Research Shallow Water Acoustic Clutter Program. The main objectives of the Program are to (1)

determine the spatial and temporal variability of clutter in long range active sonar in continental shelf environments with generally low bathymetric relief, (2) identify the dominant sources of clutter, (3) understand the physical mechanisms that lead to these prominent returns, and (4) analyze the bistatic scattering characteristics of these dominant sources. In this context, clutter refers to scattering from objects in the environment that stand prominently above the diffuse and temporally decaying reverberation background that can camouflage or be confused with the returns from an intended target.

Another purpose of ACRE was to test the possibility of using low-frequency active sonar systems to rapidly image subseafloor geomorphology over wide areas in shallow water. In deep water, several remote sensing experiments have consistently and repeatedly observed strong and deterministic clutter return from lineated ridges and scarps on seamounts.^{8–11} It was found that the level of the clutter is

^{a)}Now at Northeastern University, Boston.

^{b)}Electronic mail: purnima@ece.neu.edu

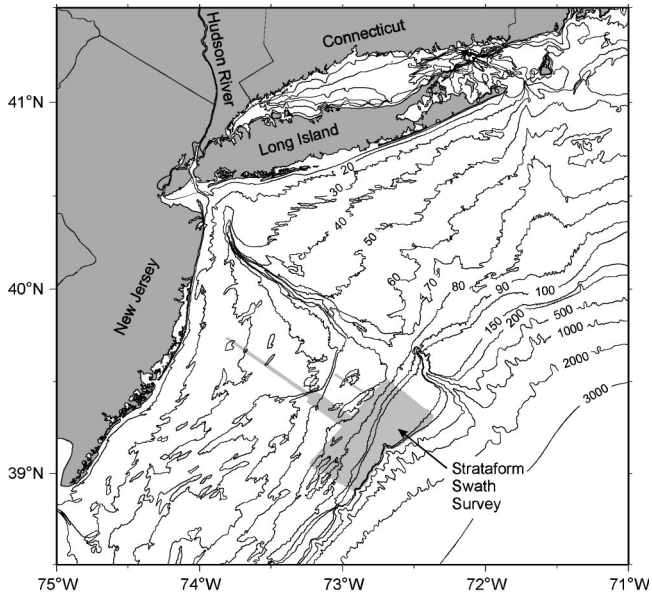


FIG. 1. Location of the East Coast STRATAFORM area off the New Jersey coast. Contours in meters.

proportional to the projected area of the scarp along the path from target to bistatic source and receiver.^{8–10} Attempting to draw an analogy with the deep water results, a number of investigators, for example in Ref. 12, proposed subseafloor geomorphology, which exists throughout the continental shelf, to be potential source for anomalous sonar returns in areas of level bathymetry in shallow water.

The New Jersey STRATAFORM¹³ site, shown in Fig. 1, was well-suited for ACRE because a number of substantial geophysical surveys^{14–16} have previously characterized seafloor and subbottom features over wide areas. Furthermore, several other acoustic experiments have been conducted in this area to investigate acoustic propagation conditions¹⁷ and to invert for seabottom properties.^{18,19} Data from the geophysical surveys are used here to identify natural features of the seafloor and subseafloor that might possibly be imaged by our remote acoustic sensing system.

During the ACRE, two research vessels were used to acquire both monostatic and bistatic scattering data. Roughly 3000 waveforms were transmitted into the water column from vertical source arrays in the 390- to 440-Hz band and received by a horizontal towed array. A wide-area acoustic intensity map (image) of the environment was generated for each transmission. On average from 10 to 100 discrete and localized scattering events were registered per transmission. From a clutter perspective, this gives a total of at least 30 000 scattering events that could be confused with that from a large submerged vehicle over the period of the experiment. Two acoustic targets,^{1,20} essentially cylindrical air-filled elastic tubes approximately 30 m long, with known scattering properties and locations were vertically deployed and used to calibrate returns and confirm theories about waveguide scattering from extended objects and long-range imaging in continental shelf environments.

In this paper, wide-area acoustic images from the ACRE are presented to illustrate the prominent and discrete scattering events measured in the New Jersey continental shelf en-

vironment. Returns that coincide with the location of the calibrated targets are evident in the images indicating our ability to accurately chart the returns in both space and time. Most of the charted clutter were found to occur in regions where the bathymetry is locally level and do not coregister with any known geologic features of the seafloor. Some of the charted clutter appears to occasionally correspond with buried river channels identified from geophysical surveys. However, statistical analysis of the clutter in regions where the subbottom geomorphology has been mapped shows that there is no significant difference between the frequency of occurrence per unit area of repeatable clutter events that chart within areas occupied by buried river channels and those that chart outside of areas occupied by channels.

In Sec. II, we provide a description of the New Jersey STRATAFORM area geology and a detailed description of the ACRE. In Sec. III, we explain how long-range acoustic data are processed to generate wide-area images. Images showing the charted clutter and returns from the calibrated targets are presented in Sec. IV. Possible mechanisms for the clutter returns are discussed in Sec. V. Oceanographic data collected during the experiment, such as sound speed structure, are presented in the appendix.

II. DESIGN AND IMPLEMENTATION OF THE ACRE

Prior to the ACRE, a number of geophysical surveys^{14–16} at the New Jersey STRATAFORM site characterized the seafloor and subbottom features over wide areas. Figure 2 shows the water depth at the STRATAFORM site where bathymetric data are available at 30 m horizontal resolution.¹⁵ Seabed and subbottom features identified from the geophysical surveys are overlain on the bathymetry. The candidate features for prominent scattering include incised or buried river channels, relict iceberg scours and surface erosional features on the seafloor, and surface or near surface outcroppings of seismically reflective subsurface strata within the seabed. Figure 3 shows an interpretation of a seismic profile from Ref. 14 of several river channels buried at different depths from the seafloor intersecting with various sub-bottom strata. Apart from these geologic features, aggregates of compact scatterers such as gravel deposits on the seafloor, gas pockets in the seabed, and large and densely populated schools of fish are also possible sources of clutter.

Figure 4 shows the directional derivative (DD) of bathymetry at the STRATAFORM area. The DD is defined in Ref. 21 as the dot product of the gradient of bathymetry with the horizontal unit vector pointing in the direction of an observation point. In Fig. 4, the DD is plotted for an observation point to the north of the STRATAFORM area. Seafloor surfaces facing the observation point have positive DD, surfaces facing away have negative DD, and level surfaces have zero DD. Inspection of Fig. 4 shows that the experiment site has mostly benign slopes of less than 0.5° with few discrete features on the seafloor. Even the seafloor features that are noticeable, such as the iceberg scours and erosion pits, have small slopes typically $<3^\circ$. Their vertical relief is typically smaller than the acoustic wavelength of roughly 4 m in this experiment.

In both Figs. 2 and 4, the subbottom features shown are based on interpretations of geophysical data acquired *prior* to the ACRE 2001. The areas where the subsurface features are shown are the only areas where the subbottom data had been acquired and analyzed prior to the ACRE 2001. There may be other subsurface features elsewhere in the figure, but the geophysical data had either not been collected in those areas or had not been analyzed prior to the experiment. The most current geophysical interpretations of the subsurface features is provided and used in our analysis of the measured clutter in Sec. IV.

The experiment was conducted using the Research Vessel (RV) *Endeavor* and the NATO Research Vessel (NRV) *Alliance* (Fig. 5). RV *Endeavor* was used mainly as a source ship for bistatic measurements. It deployed a source system used by the Multistatic ASW Capability Enhancement Program (MACE) consisting of a seven-element array of uniformly spaced XF-4 transducers, beamed to transmit at broadside during the whole experiment. RV *Endeavor* was fastened to moorings at three sites where a fixed transmission location was maintained for the bistatic measurements. NRV *Alliance* was the only ship that deployed a horizontal receiving array. It also deployed a two-element MOD40 transducer source system that was towed for quasi-monostatic measurements. A nominal tow speed of 2 m/s was maintained throughout the measurements by NRV *Alliance*. The horizontal receiving array was a 256-element line array with three nested apertures, each consisting of 128 sensor elements evenly spaced at 0.5, 1, or 2 m, respectively. Only data from 128 elements at 1-m spacing are analyzed in this paper. For a sound speed of 1500 m/s, this subaperture corresponds to an array cut for 750 Hz. Each of the two calibrated targets,^{1,20} deployed at a selected site, were moored to the bottom in waters approximately 80 m deep, 18 m off the seafloor. They stood vertically in the water column spanning roughly 32- to 62-m depth from their own buoyancy.

The tracks traversed by NRV *Alliance*, mooring locations of RV *Endeavor*, and locations of the two calibrated targets and subbottom features are also plotted in Fig. 4. Acoustic transmissions were centered about three distinct sites of the STRATAFORM area identified as sites 1, 2, and 3. The water depth at the three sites ranged from approximately 70 to 130 m. At site 1, the tracks of NRV *Alliance* and the source location for RV *Endeavor* were designed to image the buried river channels and shallow subsurface reflectors truncated at the seafloor.^{14,15} At site 2, the erosion pits on the seafloor and the subsurface reflectors, as well as the calibrated targets, were the focus of the measurement. The tracks at site 3 were designed to image the iceberg scours and more subsurface reflectors and to obtain scattered returns from the Hudson Canyon walls. The large number of tracks at each site was necessary to study the range and azimuth dependence of the scattering and to distinguish returns from the various candidate clutter targets at a variety of ranges and azimuths. The numerous tracks also served to help break the ambiguity inherent in line array measurements.

The sources transmitted both linear frequency modulated (LFM) and sinusoids or “continuous wave” (CW) signals of varying duration in the frequency range from 390 to

440 Hz.¹ For the bistatic transmissions by RV *Endeavor*, the LFM signals were shaded with a Tukey window while the CW signals were shaded with a Hann window. For the quasi-monostatic transmissions by RV *Alliance*, a rectangular window was used in all transmissions for both the LFM and CW signals. The length of each NRV *Alliance* track line is roughly 10 km and the waveforms were transmitted at every 50- or 100-s interval. With a speed of 2 m/s for the receiver ship, data from a total of roughly 50 or 100 transmissions were measured along each track.

III. WIDE-AREA IMAGES OF THE OCEAN ENVIRONMENT

A. Generating images in near-real-time

During the ACRE, a wide-area image of received sound pressure level as a function of horizontal position over tens of kilometers was generated for every transmission (ping) in near-real-time. For a given transmission, two-way travel time was used to determine the range of returns and beamforming to determine the azimuth. The process has been previously described in Refs. 8, 21, and 11. It follows the same principles used in high-frequency side-scan sonar, medical ultrasound, and radar image processing except that the present imaging process involves the complexities of multipath propagation, waveguide scattering, and dispersion.

In this paper, echo returns from mono- and bistatic LFM transmissions of varying duration measured with NRV *Alliance*'s horizontal line array with 128 elements at 1-m spacing are analyzed. The raw time series data for each hydrophone were filtered, demodulated, and decimated. This decimated array data were then converted to beam-time data by time-domain beamforming. The sharpest cross-range resolution is at broadside, where the signals arrive almost perpendicular to the array axis. The broadest cross-range resolution is near endfire, where the signals arrive almost parallel to the array axis.

A Hanning spatial window function was applied in the beamforming to reduce sidelobe levels where the first sidelobe level is down 30 dB from the main lobe. This was important for reducing sidelobe leakage from radiated noise present in beams containing the source ship which was moored near the features to be imaged. Following Eqs. (1) and (2) of Ref. 21, the 3-dB beamwidth β of the array is approximated using

$$\beta(\varphi) \approx 1.3 \frac{\lambda}{L \cos \varphi}, \quad (1)$$

for steering angles from broadside $\varphi=0$ to a transition angle φ_t near endfire, $\varphi=\pi/2$, where λ is the wavelength and L is the array length. As φ approaches φ_t , ambiguous beamwidths each reach a value approximately equal to that at endfire, and begin to merge until they completely overlap at endfire where the beamwidth is

$$\beta(\varphi = \pi/2) \approx 2.8 \sqrt{\lambda/L}. \quad (2)$$

For the array aperture of 127 m, the optimal 3-dB resolution is about 2.1° at broadside and 27° at endfire. This corresponds to a cross-range resolution $\Delta\rho=\rho\beta$ of 364 m at

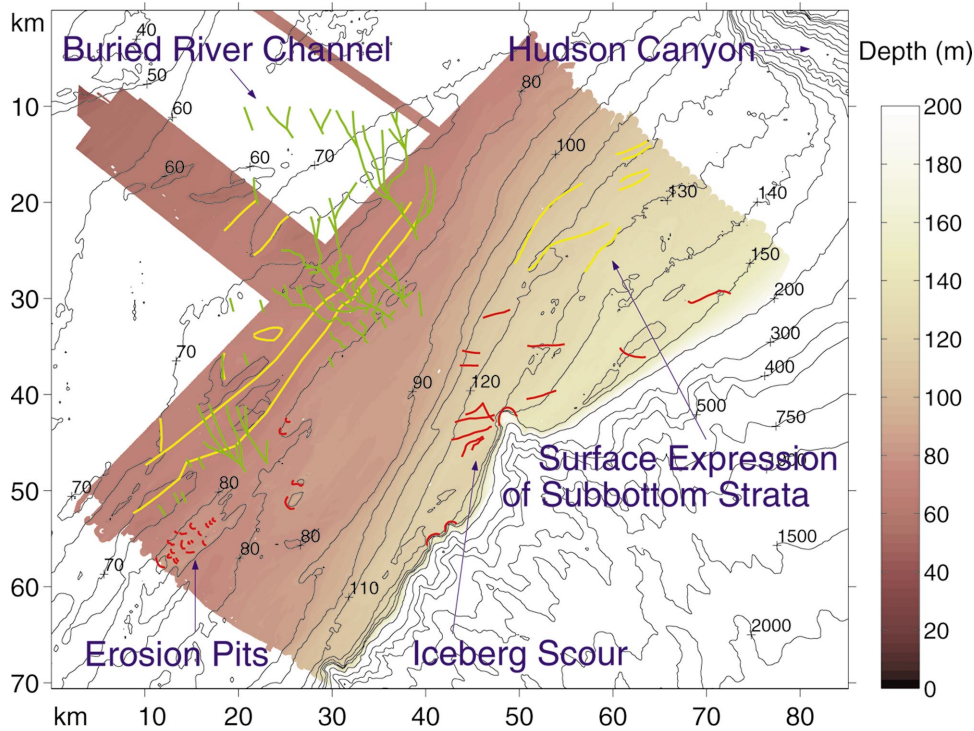


FIG. 2. Bathymetry of the STRATAFORM area with shaded region sampled at 30-m interval. Candidate features identified from previous geophysical surveys^{14,15} that might give prominent and coherent scattered returns include incised or buried river channels (green), relict iceberg scours and erosion pits on the seafloor (red), and surface or near-surface expression of seismically reflective subsurface strata within the seabed (yellow). Coordinates of x and y axis origin in the north-west corner: $39^{\circ} 31.00'N$, $73^{\circ} 17.28'W$.

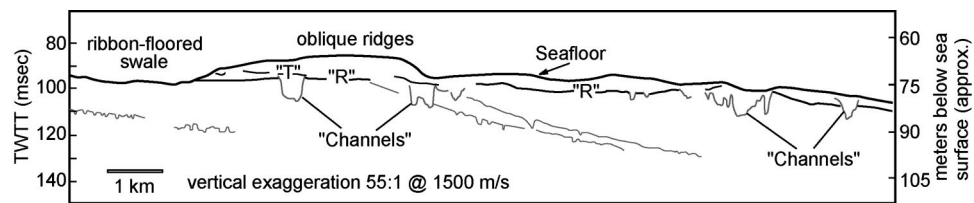


FIG. 3. Interpreted seismic line, adapted from Ref. 14, showing the shallow stratigraphy within site 1. The figure shows numerous buried river channels intersected by highly reflective subbottom strata. One of these subbottom strata is called an R-reflector.

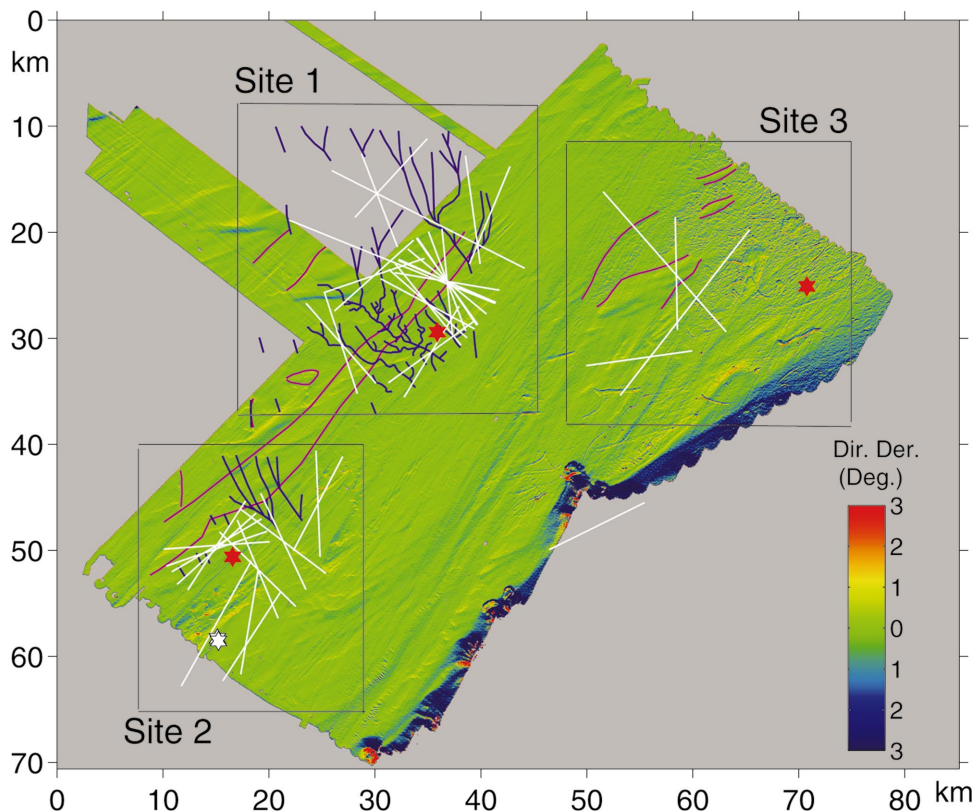


FIG. 4. Directional derivative of the 30-m sampled bathymetry at the STRATAFORM site with respect to a source far in the north. The seafloor is mostly level locally with slopes of $<1/2^{\circ}$. There are very few discrete features such as iceberg scours and erosion pits on the seafloor with slopes of at most 3° . Acoustic transmissions were centered about three distinct sites in the STRATAFORM area. Overlain are the tracks traversed by NRV Alliance (white lines), mooring locations of RV Endeavor (red stars), location of the two calibrated targets (white stars), and subbottom features (blue and pink lines). Coordinates of x and y axis origin in the north-west corner: $39^{\circ} 31.00'N$, $73^{\circ} 17.28'W$.

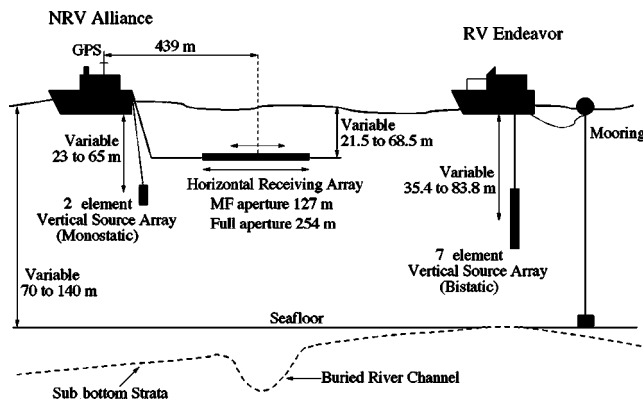


FIG. 5. A sketch of NRV Alliance towing a two-element vertical source array at 1.875-m spacing and a 256-element horizontal receiving array with spacing between elements of 0.5, 1, and 2 m. Only data from 128 elements of the receiving array at 1-m spacing are analyzed in this paper. The mean depth of the NRV Alliance two-element source was varied between 23 and 65 m. The mean depth of the receiver array was varied from 22 to 68 m. RV Endeavor was moored at each site where measurements were collected and it deployed the MACE source system consisting of a 7-element array spaced at 1.625 m. The mean depth of the MACE source was varied between 35 and 84 m and it was beamed to transmit at broadside throughout the experiment. The water depth at the three measurement sites ranged from approximately 70 to 130 m.

broadside and of 4.7 km at endfire at a range of $\rho=10$ km, which is a typical range for detecting the clutter events during the experiment.

The beam-time data are linearly converted to beam-range data by multiplying the total two-way travel time with half the mean sound speed c which was taken to be 1475 m/s based on measured sound speed in the water column (see Appendix A). To improve on the range resolution and signal-to-additive-noise ratio, the LFM data from 390 to 440 Hz were match filtered with a replica of the source signal to give an effective range resolution of $c/2B \approx 15$ m, where the bandwidth B is 50 Hz. The data were then averaged to 30 m resolution and then mapped to a Cartesian grid with the same 30×30 m² grid increment as the high resolution bathymetry data used. The mapping procedure accounts for beam overlap by an incoherent averaging of adjacent beams.²²

Multi-modal propagation in a shallow water waveguide leads to time spread and delay in the mean arrival time of signals. This dispersion effect is a result of the differing modes of the waveguide propagating the acoustic energy with varying group velocities, and is highly dependent upon the waveguide properties. It also varies as a function of range and depth due to modal interference which causes a change in the spatial and temporal structure of the signal. Simulations in various shallow water waveguides in Ref. 23 show that the error in charting returns from objects in the water column is about 200 m for the ranges relevant to this study after match filtering with the source signal. Charting error varies with the sound speed used to convert travel time to range. The charting error of 200 m is for a sound speed that corresponds to the minimum sound speed in the water column. This sound speed was found to provide the smallest errors in localizing a source or target in the water column.

During the experiment, fluctuations on the order of $\pm 3^\circ$ were observed in the towed array heading sensor measure-

ments. At site 2, the GPS positioning of the calibrated targets relative to the source and receiver was used to provide a more accurate mean orientation of the receiver array for each track. The corrections needed for each track, however, were small, approximately 0° to 4° . Since no calibrated targets were present at sites 1 and 3, no corrections were applied to those images.

The standard deviation of a pixel value or an average of stationary pixel values in an acoustic image is now estimated.⁸ We assume that the transmitted waveform's interaction with the seafloor scattering area completely randomizes the return such that the real and imaginary temporal components of the instantaneous scattered field are identically distributed and uncorrelated zero-mean Gaussian random variables. The instantaneous intensity of the return is then exponentially distributed and the time-averaged intensity is gamma distributed²⁴ with degrees of freedom μ corresponding to the time-bandwidth product TB of the scattered field,²⁵ where T is the measurement time. This product is an approximate measure of the number of independent and instantaneous intensity fluctuations averaged over T . If the reverberation level in dB *re* 1 μ Pa for a given pixel in the acoustic image is $R=10 \log$ (mean square pressure), the standard deviation σ of the reverberation level in dB is^{8,25}

$$\sigma(\mu) = 10(\log e) \sqrt{\zeta(2, \mu)}, \quad (3)$$

where

$$\zeta(\nu, \mu) = \sum_{k=0}^{\infty} \frac{1}{(\mu+k)^\nu}, \quad \text{for } \nu > 1, \quad \mu \neq 0, -1, -2, -3, \dots, \quad (4)$$

is the Riemann's zeta function. For an instantaneous intensity measurement, $\mu=1$ and $\sigma(1) \approx 5.6$ dB. To reduce the standard deviation to 3 dB which is relatively negligible compared to the dynamic range of the levels spanned in the images produced in this experiment, an averaging time of $T=0.04$ s is used. This corresponds to $\mu=2.0$ for the $B=50$ Hz bandwidth LFM transmissions analyzed in this paper. Stationary averaging of adjacent pixels would lead to a smaller standard deviation according to Eqs. (3) and (4), but will also reduce the spatial resolution. (We ignore averaging of overlapping beams because the measurements are not independent.) The prominent returns typically stand above the background reverberation by tens of decibels (many standard deviations) and are therefore considered to be deterministic.

B. Interpreting images generated with low frequency and long range towed array sonar

Figure 6 shows a wide-area image of the ocean environment at site 2 obtained from a single bistatic transmission with the receiver ship NRV Alliance oriented along track 17. The origin of the range axis along the x and y directions in the figure is at $39^\circ 31.00'N$, $73^\circ 17.28'W$. This origin is at the north-west corner of the images in Figs. 2 and 4. This fixed location for the range origin will be used for all acoustic intensity images presented in this paper. The array heading is measured with respect to true north increasing in a clockwise direction. The image shows both the diffuse back-

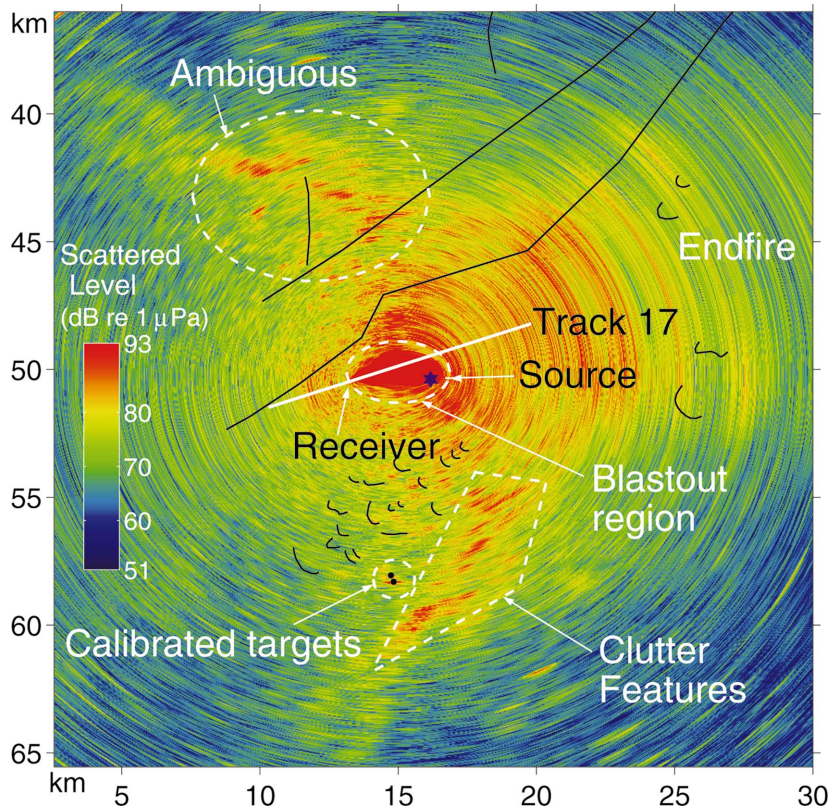


FIG. 6. A single-ping bistatic wide-area image along track 17 at site 2. Travel time to range conversions are done by multiplying the two-way travel time with half the mean sound speed of 1475 m/s. All returns are mirrored about the array axis (71° with respect to true north) due to left-right ambiguity. Two prominent and discrete scattering events >20 dB above the background co-register well with the location of the calibrated targets approximately 8.5 km to the south. Numerous other prominent scattering events (clutter features) that can be confused with returns from the calibrated targets are present in the image. Comparison with Fig. 8 breaks the receiver line array's left-right ambiguity and places the true location of these clutter features to the south within the dotted white trapezoid. (Date: 1 May, ping time: 13:04:30Z, transmission: 1-s duration LFM from 390 to 440 Hz, mean source depth: 55 m, mean receiver depth: 42 m, array heading: 71° .)

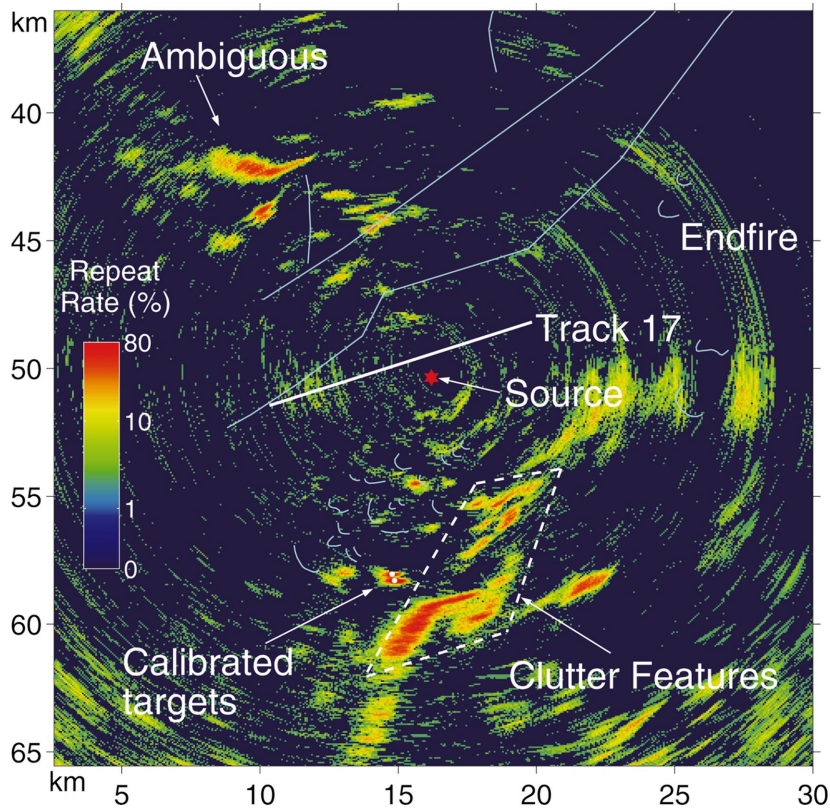


FIG. 7. Hotspot consistency chart from 49 bistatic LFM transmissions along track 17. The figure displays the number of images that register a strong scattering event >10 dB above the local average within 1.8×1.8 km² area of a given pixel location. Scattering events within the trapezoid as well as those from the calibrated targets are consistently prominent in most of the wide-area single ping images along track 17. Events within the trapezoid can either be confused with or camouflage those from the calibrated targets. (Date: 1 May, time: 12:37:00Z to 14:01:00Z, mean source depth: 55 m, mean receiver depth: 42 m, array heading: 71° .)

ground reverberation level in decibels from the ocean environment as well as strong scattered returns from the calibrated targets and other submerged objects or features. For each transmission, the signal first measured by the receiving array is that arriving directly from the source. This direct arrival is strong and gives rise to the blast-out region which appears as a red ellipse surrounding the source and the receiver in Fig. 6. The diffuse background reverberation scattered from random rough patches of the ocean environment, as well as the returns from the calibrated targets and other submerged objects or features arrive after the direct signal has passed. This environmental reverberation has a mean intensity that decays with range due to spreading and absorption loss in the waveguide. The rate of decay depends on the properties of the waveguide, such as the sound speed structure, attenuation in the water column and bottom, and surface and bottom inhomogeneities, as well as the measurement geometry. The decay in the reverberation provides vital information about the environment needed to model propagation and diffuse scattering. In Fig. 6, we do not average out the trend in the data since it would eliminate this vital information.

A horizontal line array has left–right ambiguity about the array axis that is expressed differently in monostatic and bistatic charts. Prominent returns are ambiguously charted nearly symmetrically about the receiving array axis in monostatic geometries. For bistatic geometry, ambiguity occurs on an ellipse with a major axis that passes through the source and receiver. A diagram illustrating the two-way travel time ellipse for bistatic measurements is provided in Ref. 9, where it is shown that distortion in the image may occur as the ambiguous returns are charted to either a smaller or broader spatial extent. The methods used in this paper to resolve the left–right ambiguity in the measurements are discussed further in Sec. III D. In Fig. 6, the left–right ambiguity is mapped onto ellipses at close ranges to the bistatic source and receiving array. At ranges larger than the source–receiver separation, the ambiguity about the array axis approaches circular symmetry.

During the ACRE, few monostatic measurements were made in comparison to the bistatic measurements because the monostatic source was weaker. As a result, we were not able to image very far out in range with the monostatic source. The rough maximum range before the acoustic imaging system became noise limited was on the order of 10 km for the monostatic source. For the bistatic source, this range was more than 50 km away from the source. Most of the images illustrated in this paper are from bistatic measurements.

C. Hotspot repeatability chart

In order to measure the frequency of occurrence of a strong scattering event from a target or feature along a given track, we generate a hotspot consistency chart for each track. First, a moving local peak detector is applied to a single wide-area image to detect pixels where the reverberation level stands more than 10 dB above the average for a $1.8 \times 1.8 \text{ km}^2$ subimage. These pixels are assigned a value of 1 and all others 0. Pixels in the blast-out region are assigned a

value of 0, since we are only interested in detecting targets and features in the environment. This binary matrix from each reverberation image is then summed for all the images along a track to form the hotspot consistency chart for each track. The hotspot repeatability chart for track 17 is shown in Fig. 7. A total of 49 bistatic acoustic images are combined to form this composite. In Fig. 7, a given pixel shows the number of images that register a strong scattering event $>10 \text{ dB}$ above the local average within $1.8 \times 1.8 \text{ km}^2$ area of the pixel. The maximum of the legend in each hotspot repeatability chart is the maximum clutter repeatability in the given image.

The standard deviation²⁵ for the time-averaged diffuse reverberation is 3 dB as shown in Sec. III A. The local peak detector algorithm thus picks out pixels with levels that are about three standard deviations above the local mean within each subimage. Both scattered returns from fixed targets and diffuse background reverberation decay with range due to spreading and absorption loss. A globally fixed threshold detector is only useful if the decaying trend of the reverberation can be accurately removed before the detector is applied. It is usually difficult to detrend measured data accurately given the lack of *a priori* environmental information needed to model the trend. Our local peak detector takes the range decay into account empirically. It is also independent of the strength of the source, making it versatile and easy to implement.

We combine information contained in wide-area images acquired along a single track to form the hotspot consistency chart for the given track. This chart is found to be extremely useful in picking out strong and persistent echos as well as in resolving ambiguity in measurements along a single track as will be discussed in Secs. III D and IV. A hotspot consistency chart combining information from multiple track lines should be used with caution. Our attempts at combining data from all the tracks at a given site by forming an overall hotspot chart for the site were found to lead to confusion in identifying actual returns. This is because the many tracks at a given site have varying array headings and the same scatterers often do not appear on different tracks, making it impossible to resolve ambiguity by such a combination. Even when the same scatterers do appear on differing tracks the combination does not significantly reduce ambiguity but rather leads to multiple ambiguities that often mask true features as noted in Ref. 22.

D. Ambiguity resolution

Two methods are employed in the present analysis to resolve the left–right ambiguity in the returns from the horizontal line array. The first method involves comparing events from images obtained on tracks that have approximately similar location but where the array orientation differs.⁸ In these tracks, the true returns will be consistently charted to the same location, while the ambiguous ones will be charted to different locations.

The second method resolves the ambiguity for measurements made along a single track with the use of the hotspot consistency chart for the track. This method is applicable in bistatic scenarios when the range to the feature is not much

larger than the source–receiver separation. It exploits the elliptical shape of the travel time locus to break the ambiguity in the line array measurement.

These two methods for ambiguity resolution will be pointed out in the images in Sec. IV.

IV. LONG RANGE ACOUSTIC IMAGING RESULTS

In this section we present the major experimental findings of the ACRE 2001. Roughly 3000 waveforms were transmitted into the water column and wide-area image of the reverberation from the ocean environment was generated in near-real-time for each transmission. From 10 to more than 100 discrete and localized clutter events (10 dB above the background) were registered for each image. This gives a total of at least 30 000 scattering events that could be confused with that from a large submerged vehicle over the period of the experiment. The vast majority of prominent and discrete scattering events appear in areas where the bathymetry is locally flat or with slopes less than 0.5° . This implies that these returns most probably do not originate from the seafloor and that they could be due to objects in the water column or possibly features in the subbottom.

In the following three subsections, individual ping and hotspot consistency images from various tracks at the three sites are shown. The clutter in these images is representative of the scattering events measured at each site. We begin our discussion with site 2 where the calibrated targets were deployed.

A. Site 2 with calibrated targets

In this section, we present wide-area images showing strong and discrete returns from the calibrated targets deployed in the waveguide at site 2. These images, which were acquired during the middle of the experiment, also show an organized pattern of prominent scattering events in a region of effectively flat bathymetry (slopes less than 0.5°), with water depths from 80 to 100 m, where the subbottom had not been well profiled. Subsequent geophysical surveys, specifically designed to explore the subbottom in the vicinity of these prominent scattering events at site 2, discovered a network of buried river channels.^{26,27} Most of the prominent scattering events, however, did not coregister with the newly discovered river channels. The measurements at site 2 were repeated about 4 days later at the end of the experiment along tracks with nearly similar range and bistatic location for the track center, but with varying orientation. These measurements found the clutter events at site 2 to be highly variable across the tracks, all of which had differing headings. Since there was a 2-h time period for the data collection on each track, the clutter could also have evolved with time. A statistical analysis of the spatial distribution of repeatable clutter is given in Sec. IV A 2 to determine if the repeatable clutter favors any particular spatial location.

In Fig. 6, which shows a wide-area image from track 17 at site 2, we observe two prominent scattering events approximately 8.5 km away to the south of the source and receiving array that register well with the calibrated targets. These events stand out by more than 20 dB above the reverberation in surrounding areas. This difference is much larger

than the received level standard deviation of 3 dB (shown in Sec. III A). The receiving line array has left–right ambiguity about the array axis. We therefore see the ambiguous returns from the calibrated targets charted to the west in Fig. 6.

We observe numerous other prominent and discrete scattering events in the image shown in Fig. 6. Many of these features are just as prominent as the calibrated targets, standing out by more than 20 dB above the reverberation in surrounding areas. These features can be confused with or camouflage the calibrated targets if the precise location of the targets are unknown.

To identify where the clutter originates, we break the receiving array's ambiguity by comparing Fig. 6 with Fig. 8, a bistatic transmission along track 14 which is at a similar range but slightly different orientation than track 17. In both Figs. 6 and 8, we observe consistently strong and prominent events within the trapezoid to the south of the source and receiving array. This shows that the region with the trapezoid is the true location of the clutter. The ambiguous events are charted to the west in both figures and are not in the same location. The differing array orientation in tracks 14 and 17 causes the false returns of the features to be charted to different regions.

To examine the repeatability of the prominent scattering events within the trapezoid as well as those from the calibrated targets, a hotspot consistency chart, as described in Sec. III C, is derived for track 17 and shown in Fig. 7. From Fig. 7, we observe that scattering events from the calibrated targets are repeatedly prominent as the receiver ship moves along the track. They appear as a local maxima in most of the 49 images along the track. The scattering events within the trapezoid in Fig. 7 are also prominent and repeatable. They appear in most of the charts with levels >10 dB above the local average.

In Fig. 9, we overlay the prominent events from Fig. 7 on the color directional derivative of bathymetry. Only events with levels that are 10 dB above the local average and are consistent in at least ten transmissions along the track are overlain. From Fig. 9, we infer that the prominent and consistent scattering events within the trapezoid do not originate from the seafloor surface because the seafloor in this region is level. The prominent events could have been caused by other objects submerged in the water column or features in the subbottom. Features in the subbottom may not be detectable using a conventional bathymetric depth-sounder, but can be found by subbottom profiling.

During the ACRE 2001, RV *Endeavor* explored the bottom with a hull-mounted chirp subbottom profiler in the vicinity of the prominent scattering events within the trapezoid in Figs. 6 and 7. More detailed geophysical surveys were carried out in this region using both deep-tow and hull-mounted subbottom profiling systems in August 2001^{26,27} and May 2002 under the Acoustic Clutter Program. All three surveys found a network of buried river channels with flanks that shoal close to the seafloor in the vicinity of the scattering events. An interpretation²⁷ of the morphology of these buried river channels from the geologic data are shown in Fig. 10. Figure 11 shows a deep-tow chirp profile of several buried

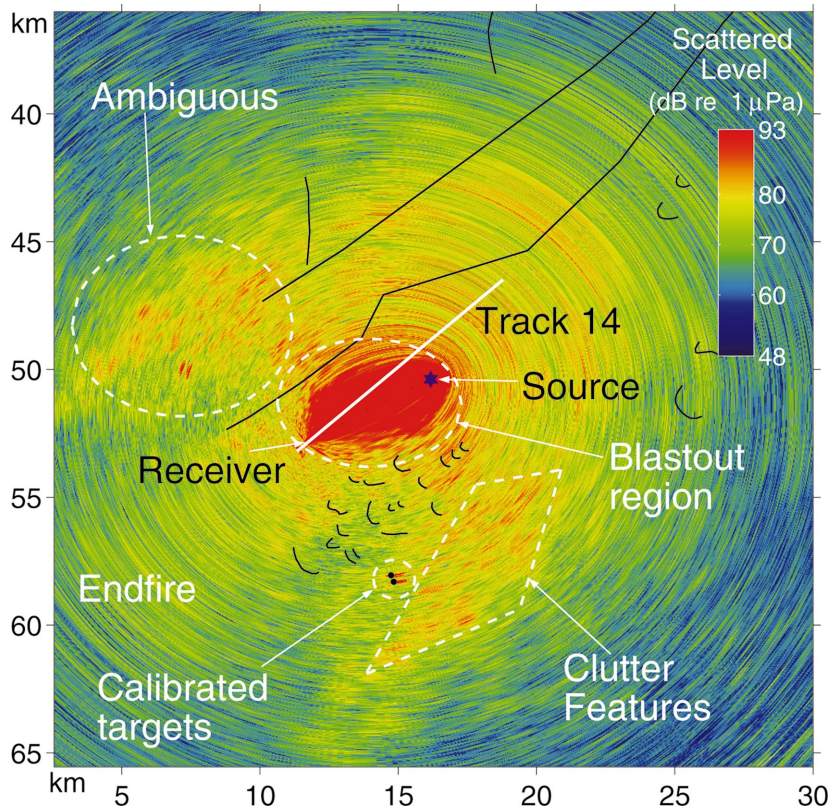


FIG. 8. Single-ping bistatic wide area image from track 14. Left–right ambiguity of the prominent scattering events from the features of interest can be resolved by comparing this figure with Fig. 6. (Date: 1 May, ping time: 11:35:00Z, transmission: 2-s duration LFM from 390 to 440 Hz, mean source depth: 55 m, mean receiver depth: 22 m, array heading: 226°.)

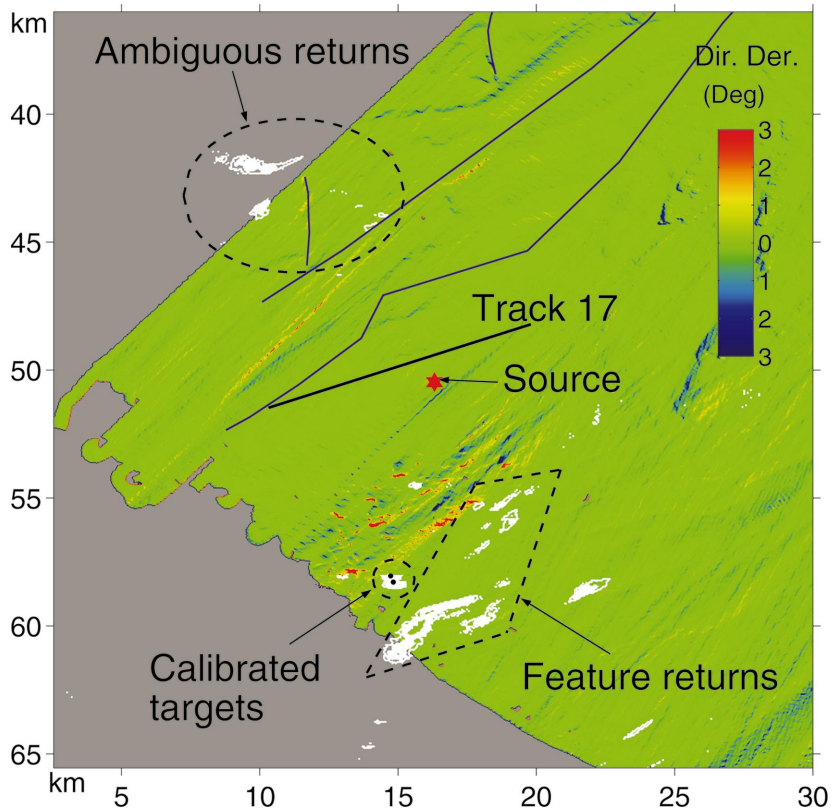


FIG. 9. Prominent events from the LFM transmissions on track 17, as shown in Fig. 7, are overlain in white on the directional derivative of the bathymetry calculated with respect to the site 2 bistatic source location and the receiver location in the middle of track 17. Only prominent (>10 dB above local average) and repeatable events that occur in at least 10 charts out of 49 are overlain. Scattering events of interest within the trapezoid do not originate from the seafloor because the seafloor in this region is level. The returns are probably due to other scatterers in the water column or in the subbottom.

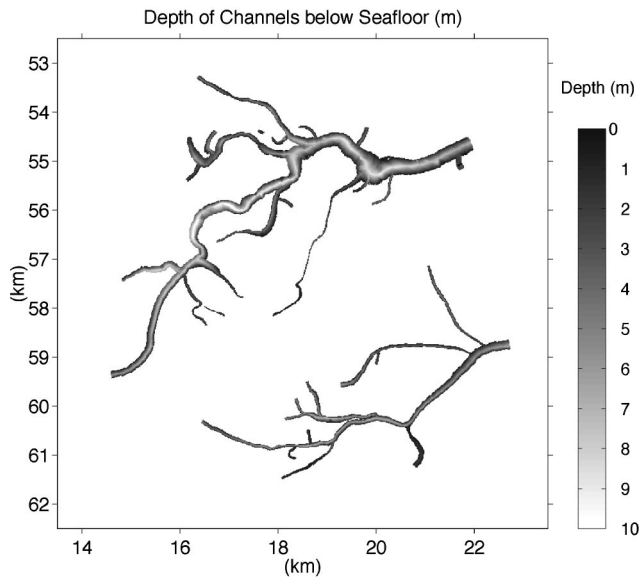


FIG. 10. Most recent interpretation²⁷ of the depth of the buried river channels at site 2 below the water–sediment interface. This interpretation is based on geophysical survey of August 2001²⁶ at site 2.

river channels found in this area from the geophysical survey in August 2001.

The trace of the buried river channels from Fig. 10 is overlain on the hotspot consistency chart for track 17 in Fig. 12. From Fig. 12, we observe that some of the prominent scattering events appear to coincide with the location of the newly discovered buried river channels. But the rest, especially the clutter in the south of the trapezoid, do not coregister. Possible source of scattering for these unidentified clutter events are discussed in Sec. V.

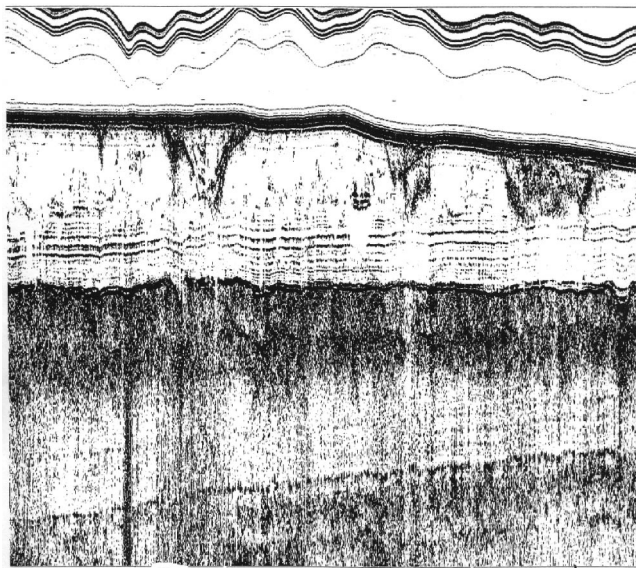


FIG. 11. Deep-tow subbottom chirp profile from the geophysical survey of August 2001^{26,27} at site 2 showing three buried river channels discovered in the vicinity of the prominent and discrete scattering events observed at this site as shown in Fig. 7. Some of these channels at site 2 shoal close to the water–sediment interface and are less deeply buried than those at site 1. These river channels are located between 39° 1.3684'N, 73° 2.9619'W and 39° 2.2878'N, 73° 2.9624'W.

1. Temporal and spatial variability of clutter

The trapezoidal region in Figs. 6–8 that registered strong scattering events on tracks 17 and 14 also registered many clutter events with significant temporal and spatial variability over the course of the experiment. In general, these clutter events were found to be consistently observable over hours, but not longer than a day.

Apart from tracks 14 and 17, scattering events in the region within the trapezoid were also registered on track 23x where the data were collected on the same day as tracks 14 and 17. Figure 13 shows the hotspot consistency image composed of bistatic LFM transmissions along track 23x. We observe scattering events within the trapezoid, as well as returns from the calibrated targets. Fewer bistatic signals were transmitted along track 23x than in tracks 17 or 14, as noted in the figure.

No high-level clutter is observed within the trapezoid during track 18, a day before track 17, as shown by the hotspot consistency chart for track 18 in Fig. 14. Track 18 has the same source depth of 55 m and same mean receiver array depth of 42 m, but a different location and orientation from track 17. The scattering events originating from the calibrated targets are prominent in Fig. 14. Only sidelobe leakage from the calibrated targets in the endfire direction appears within the trapezoid.

Figures 15(a)–(d) show data from four tracks with similar centers, but different orientation from tracks 17 and 14. (Note that Fig. 15 is displayed after Fig. 16.) These data were collected within a half-day period 4 days after tracks 17 and 14 were run. Several prominent clutter events occur in the trapezoidal region in Figs. 15(a) and (c). However, little or no clutter is observed in that region in Figs. 15(b) and (d). This figure shows that clutter events in and out of the trapezoid are highly variable functions of array orientation, time, or both. Since none of the individual tracks were repeated, the purely temporal variability cannot be isolated in this experiment. No returns can be seen from the calibrated targets in the later measurements because they deflated within days and were no longer functional.

2. Statistical analysis of clutter repeatability inside and outside of areas occupied by buried river channels at site 2

Clutter data from all site 2 tracks over the region where the subbottom has been profiled are analyzed statistically. The goal of this analysis is to compare the frequency of occurrence of repeatable clutter inside and outside of regions occupied by buried river channels. For each track, a hotspot consistency subimage covering a box of dimension 9.5×10.5 km² containing the buried river channels is used in the analysis. This region is the area within the yellow box in Fig. 12. The area occupied by the buried river channels within this box is approximately 5.8% of the total area of the box. An area surrounding the targets of width 1.7×2.2 km² is excluded from the analysis. For each track with index n , the number of pixels $T_n(10\%)$ with hotspots that are repeatable in at least 10% of the total number of images N_n along the track is computed for the box. The number of these pixels $C_n(10\%)$ that chart within buried river channels is next com-

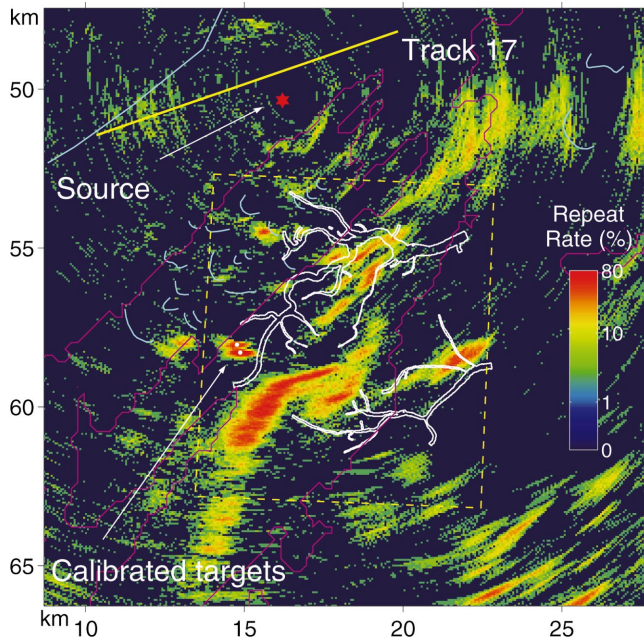


FIG. 12. Hotspot consistency chart for track 17 with white overlay showing the trace²⁷ of buried river channels discovered at site 2. This interpretation of the river channel is based on data from the August 2001 geophysical survey^{26,27} within the yellow dashed box, acquired using chirp subbottom profiling systems. Some of the prominent scattering events appear to correspond with the newly discovered river channels. Some of the events to the south, however, do not correlate with any known geologic feature in this region. This indicates that some of the prominent acoustic returns measured at site 2 are probably caused by other objects in the water column. Large and densely populated schools of fish were measured^{34,35} around the 80-m water depth contour (shown in pink line) of the STRATAFORM area and are a possible cause of some of the prominent clutter returns at this site.

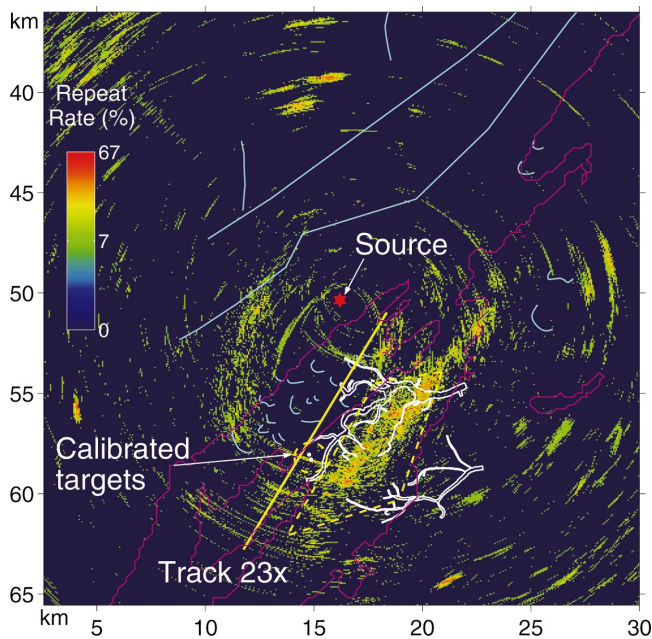


FIG. 13. Hotspot consistency chart for 15 bistatic transmissions along track 23x. Data for track 23x were collected on the same day as track 17. Prominent events >20 dB above the background are registered from the two calibrated targets approximately 8.5 km to the south of the source. Scattering events located within the trapezoid (yellow-dashed line) can be observed for track 23x, similar to that on tracks 17 in Fig. 7. The 80-m bathymetric contour is shown in pink. (Date: 1 May, time: 21:45:00Z to 23:10:00Z, mean source depth: 55 m, mean receiver depth: 43 m, array heading: 209° .)

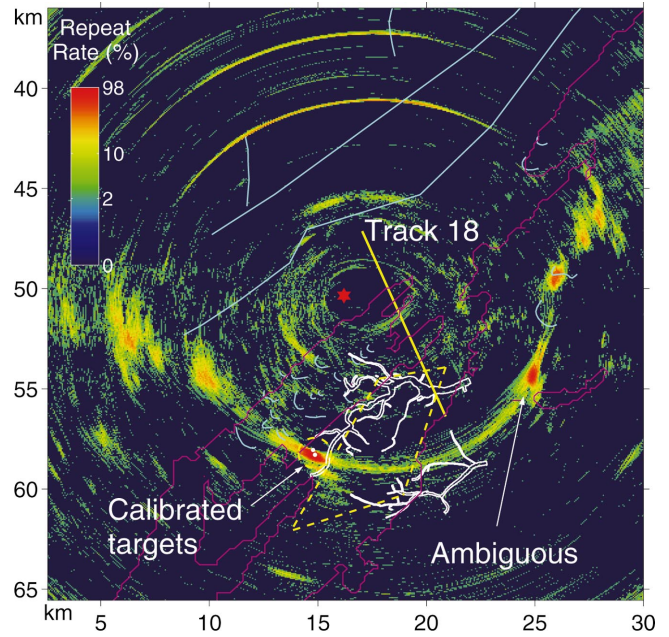


FIG. 14. Hotspot consistency chart for 47 bistatic transmissions along track 18. Data for track 18 were collected a day before that on track 17. Prominent events >20 dB above the background are registered from the two calibrated targets approximately 8.5 km to the south of the source. No high level scattering event can be detected within the trapezoid (yellow-dashed line) in this figure. The 80-m bathymetric contour is shown in pink. (Date: 30 April, time: 15:30:00Z to 16:45:00Z, mean source depth: 55 m, mean receiver depth: 42 m, array heading: 340° .)

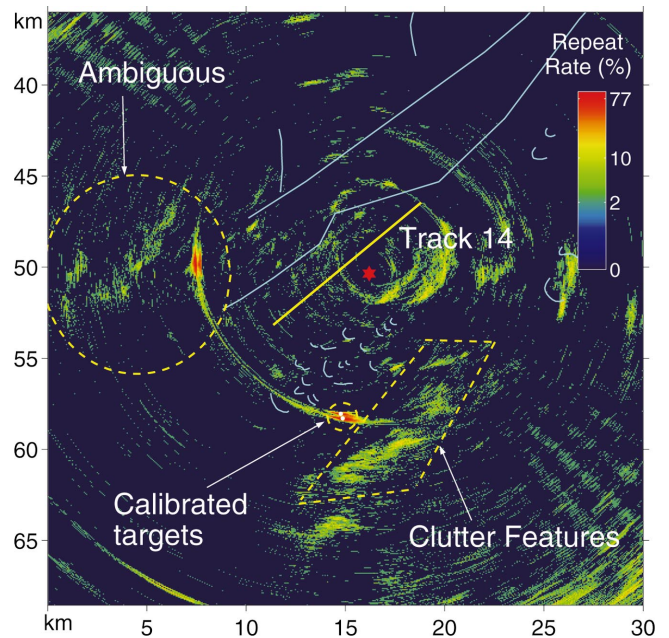


FIG. 16. Hotspot consistency image for 46 bistatic transmissions along track 14. Prominent scattering events with levels 10 dB above the background co-register with the location of the calibrated targets in most of the images. Scattering events originating from clutter features within the trapezoid are not registered as consistently along track 14. It should be noted that track 14 had a shallower receiver depth than track 17 in Fig. 7. This hotspot image shows that measurement of clutter varies with time and may also depend on receiver depth. (Date: 1 May, time: 10:30:00Z to 11:52:00Z, mean source depth: 55 m, mean receiver depth: 22 m, array heading: 226° .)

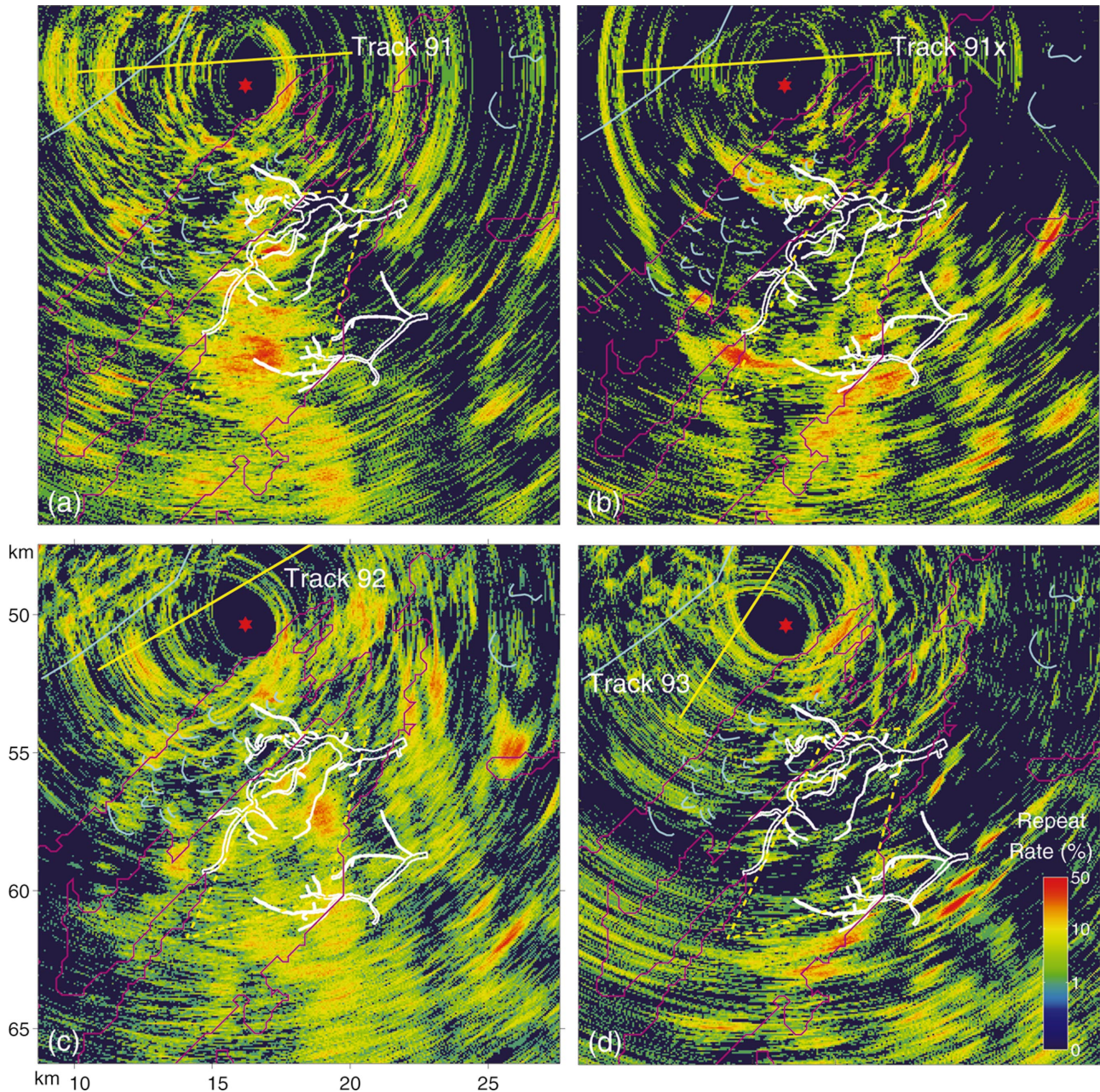


FIG. 15. Hotspot consistency charts for (a) 60 bistatic transmissions along track 91 with array heading 268.5°, (b) 42 bistatic transmissions along track 91x with array heading 270°, (c) 94 bistatic transmissions along track 92 with array heading 49°, and (d) 74 bistatic transmissions along track 93 with array heading 208.5°. Data for these tracks were collected on the same day. We observe that the scattering events within the trapezoid (yellow-dashed line) show temporal and spatial variability. The calibrated targets at site 2 had deflated during this measurement and are hence not detectable. The 80-m bathymetric contour is shown in pink. [(a)–(d) Date: 5 May, mean source depth: 35 m, mean receiver depth: 36 m.]

puted. The ratio of these quantities defined by $[C_n(10\%)/T_n(10\%)] \times 100\%$ gives the percentage $P_n(10\%)$ of repeatable clutter at 10% repeatability or more within the buried river channels. The percentage $P_n(20\%)$ is also calculated for clutter that is repeatable in at least 20% of the total number of images along the track. The results are shown in Table I. The frequency of occurrence of repeatable clutter within buried river channels has a mean from 2.3% to 4.9% and standard deviation from 1.8% to 4.3%. This frequency of occurrence of clutter within the buried river channels differs by less than a standard deviation from the area occupied by the channels in the subimage of 5.8%. This im-

plies that to within the errors of the statistical analysis, there is no evidence that the repeatable clutter favors buried river channel locations over the nonchannel locations.

3. Possible variation of clutter with receiver depth

Figure 16 shows the hotspot consistency chart for the 46 bistatic transmissions along track 14. Prominent and discrete events register at the locations of the calibrated targets in 41, or 89%, of the images. Scattering events from features within the trapezoid, however, are prominent with levels of more than 10 dB above the background in at most 7, or 15%, of

TABLE I. Frequency of occurrence of repeatable clutter within the yellow box at site 2 shown in Fig. 12. Buried river channels occupy 5.8% of the area of this box. The $M = 11$ tracks listed below are the only ones that transmitted bistatic LFM signals at site 2. Column 2 indicates N_n , the total number of images along track n , used to form the hotspot consistency chart over the box area. Columns 3 and 4 indicate $P_n(10\%)$ and $P_n(20\%)$ which are the percentages of repeatable clutter with at least 10% and 20% repeatability, respectively, that are charted within the buried river channels for each track. The unweighted mean percentages of repeatable clutter with at least 10% and 20% repeatability charted within buried river channels are calculated by taking the average of $P_n(10\%)$ and $P_n(20\%)$ respectively over all M tracks. The unweighted mean and standard deviation do not take into account the variation in the total number of images along each track. The weighted mean and standard deviation account for variations in the total number of images along each track.

Track index, n	Track name	N_n	$P_n(10\%)$ (%)	$P_n(20\%)$ (%)
1	13	46	3.2	12.0
2	14	46	2.0	0.0
3	17	49	4.4	3.2
4	18	47	4.9	10.8
5	19	9	9.3	8.2
6	20	9	4.9	3.4
7	23x	15	8.5	10.1
8	91	60	4.2	2.6
9	91x	42	3.2	3.3
10	92	94	4.5	0.0
11	93	74	1.2	0.5
Unweighted mean			4.6	4.9
Unweighted standard deviation	$\bar{P} = \frac{1}{M} \sum_{n=1}^M P_n$		3.7	3.7
	$\sigma = \sqrt{\frac{1}{M} \sum_{n=1}^M (P_n - \bar{P})^2}$			
Weighted mean			2.3	4.3
	$\bar{P}' = \frac{1}{\sum_{n=1}^M N_n} \sum_{n=1}^M N_n P_n$			
Weighted standard deviation			1.8	4.3
	$\sigma' = \sqrt{\frac{1}{\sum_{n=1}^M N_n} \sum_{n=1}^M N_n (P_n - \bar{P}')^2}$			

the images along track 14. These images were acquired when NRV *Alliance* was located towards the south-western end of the track. Clutter events measured while the NRV *Alliance* was in the north-eastern part of track 14 were less prominent within the trapezoid. Tracks 17 (shown in Fig. 7) and 14 are located at roughly the same range from the calibrated targets and the objects located within the trapezoid, but they differ in their orientation by approximately 10° . Data for these two tracks were collected on the same day, approximately 2 h apart. The receiver array was at a shallower depth in the water column of about 22 m on track 14, while it was deeper at 42 m depth on track 17. The bistatic source depth of 55 m and location did not change. These images show that variation of receiver depth may have caused substantial differences in our ability to image the clutter events within the trapezoid. Imaging of the calibrated targets is more stable with respect to changes in receiver depth. It is also highly-likely that the scatterers causing clutter may have evolved over this time leading to the observed differences in our ability to image them.

Note that the sound speed in the water column close to the location of the clutter features and calibrated targets has a maximum around 40- to 50-m depth due to a protruding filament of warm Gulf stream water as discussed in the appendix. This profile is downward refracting for a source depth of 55 m. Better detection can be obtained of scatterers located near the bottom for a deep receiver than a shallow

receiver The two calibrated targets are tall cylindrical tubes that span close to half the water column from 32- to 62-m depth which allows them to be detected much more readily regardless of the sound speed profile or receiver depth.

B. Site 3

At this site, patches of returns stretching over 35 km in range are imaged near the 100-m water depth contour on tracks 61 and 62 as seen in Figs. 17(a) and (b) respectively. Comparing Figs. 17(a) and (b) enables ambiguity about the line array to be resolved. The true location of the returns is to the south-west in both charts. It is noteworthy that the shelf break front in this region where the warmer slope water intersects the seafloor also occurs along bathymetric contour with similar water depth (see the appendix). The significance of the shelf break front to acoustic scattering is discussed further in Sec. V.

Two seafloor features are also imaged at site 3. Figures 18(a) and (b) show the hotspot consistency images formed from data acquired along tracks 61 and 62, respectively. For these particular images, levels 10 dB above the local average over an area $2.7 \times 2.7 \text{ km}^2$ are selected as peaks. The iceberg scour on the seafloor about 5 km away from the source is imaged on both tracks repeatedly. This particular iceberg scour is the only one imaged in the experiment. Scattering from the wall of the Hudson Canyon more than 30 km away from the source and receiving array is also registered as seen

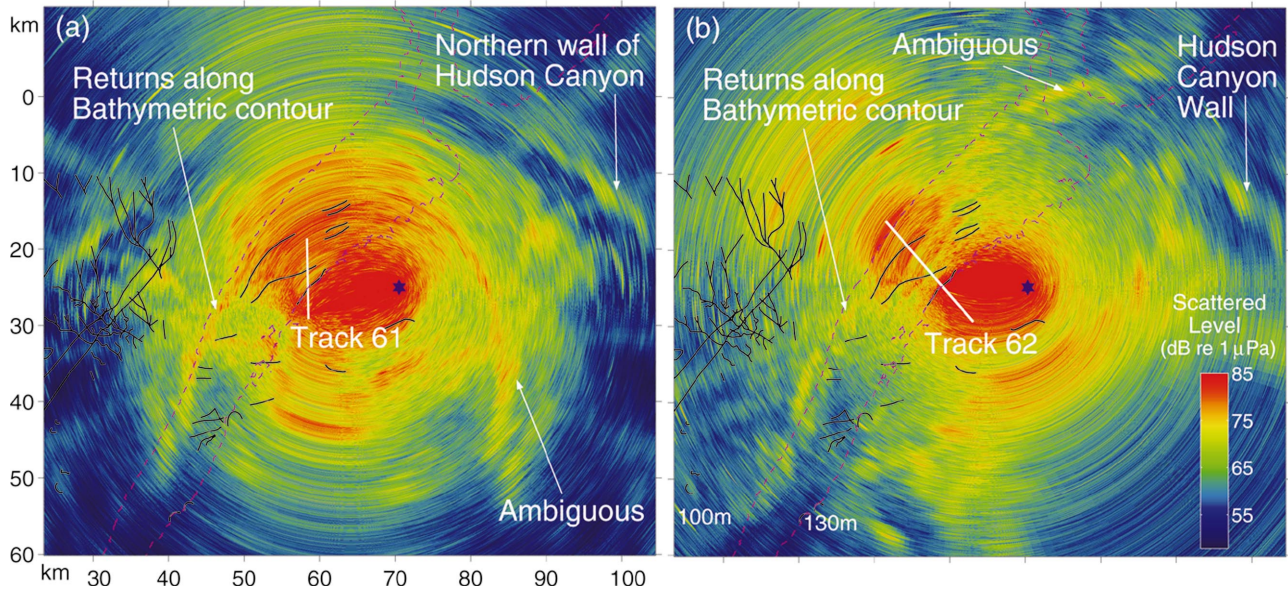


FIG. 17. Single-ping bistatic wide-area images from tracks 61 and 62 at site 3. Patches of scattered returns stretching over 30-km range and that line up with the 100-m water depth contour were imaged in this region. Comparing (a) and (b) allows the receiver line array's left-right ambiguity to be resolved. [(a): Track 61, date: 2 May, ping time: 12:11:40Z, transmission: 2-s duration LFM from 390 to 440 Hz, mean source depth: 84 m, mean receiver depth: 60 m, array heading: 358°. (b) Track 62, date: 2 May, ping time: 16:08:20, transmission: 2-s duration LFM from 390 to 440 Hz, mean source depth: 84 m, mean receiver depth: 34 m, array heading: 133°.]

in Figs. 17 and 18. Some of the patches of returns along the bathymetric contour are also visible in the hotspot consistency image in Fig. 18.

C. Site 1 with statistical analysis of clutter repeatability in and out of buried river channels

Site 1 was the focus of previous geophysical surveys where extensive networks of buried river channels were

mapped and characterized prior to the ACRE.^{14,15} Additional geophysical surveys were conducted after the ACRE to provide better characterization of the channel morphology.^{26,28}

During the first few days of the ACRE, a large number of tracks were traversed by RV *Alliance* at various ranges and azimuths from the buried river channels at site 1 shown in Fig. 19, to explore possible scattering and hence imaging of these channels. Prominent clutter events were observed

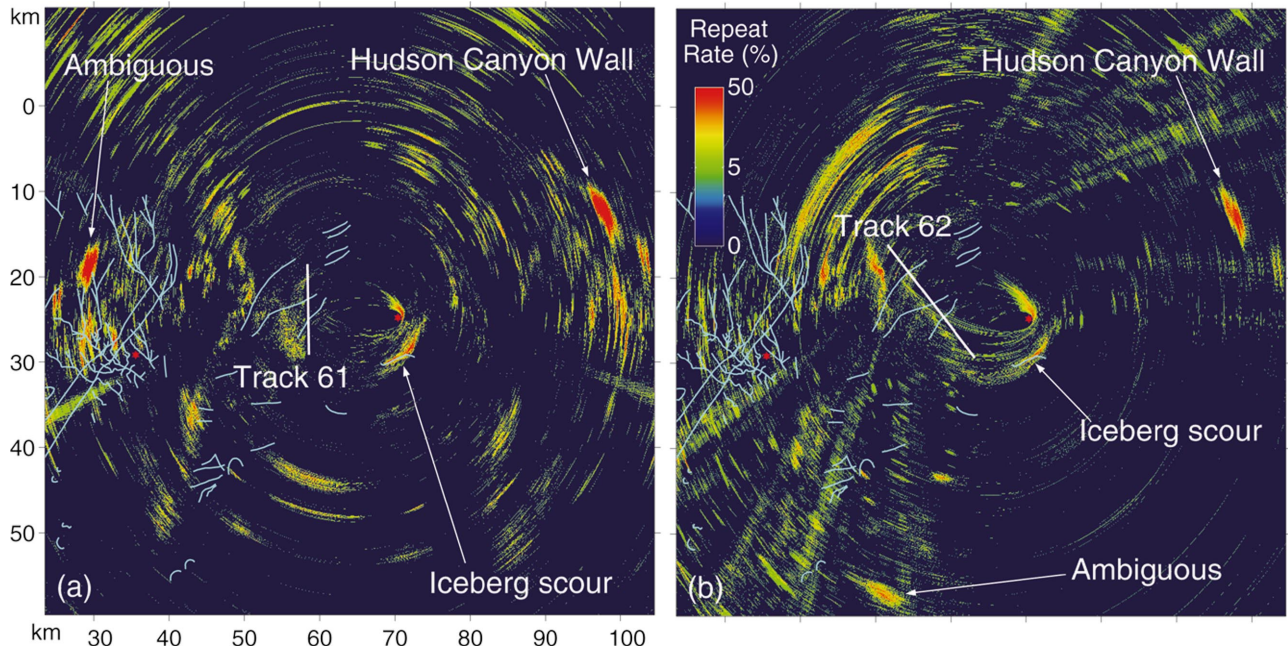


FIG. 18. Hotspot consistency charts from (a) 15 bistatic LFM transmissions along track 61, and (b) 21 bistatic transmissions along track 62 and at site 3. The figure displays the number of images that register a strong scattering event >10 dB above the local average within 2.7×2.7 km² area of a given pixel location. Comparing (a) and (b) breaks the receiver line array's right-left ambiguity. It registers prominent scattering from the walls of the Hudson Canyon located to the west in both figures, as well as scattering from the iceberg scour close to the location of the bistatic source. [(a) Track 61, date: 2 May, mean source depth: 84 m, mean receiver depth: 60 m, array heading: 358°. (b) Track 62, date: 2 May, mean source depth: 83.8 m, mean receiver depth: 34 m, array axis: 358°.]

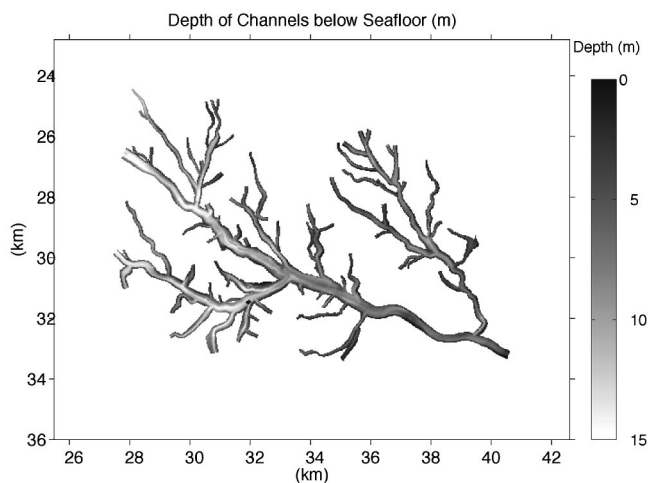


FIG. 19. Most recent interpretation²⁸ of the depth of the buried river channels at site 1 below the water–sediment interface. This interpretation is based on geophysical surveys at site 1.

throughout the region. Some of the clutter events intersect with the buried river channels. However, the registration of the clutter with the channels could not be maintained from one track to the next.

Towards the end of the ACRE, a more controlled data collection procedure was adopted where higher temporal and spatial sampling were employed to study clutter repeatability. Bistatic sonar data were collected along six tracks numbered tracks 81 to 86, located within 5-km range of the buried river channels. These tracks had the same track center but with orientations that differed by 15° from one track to the next. Rapid transmissions were made of the same LFM waveform for all these tracks. This allowed a large number of acoustic images to be captured over small variations in receiver azimuths of approximately 1°. (This angular variation is within a scatter function lobe for objects smaller than 200 m.)

Figures 20(a)–(f) show the hotspot consistency images from tracks 81–86, respectively. Figures 21(a)–(f) provide a zoomed version of Figs. 20(a)–(f) in the vicinity of the channels where the subbottom data have been collected and analyzed. In these figures, the traces of the most recent interpretation²⁸ of the river channel morphology at site 1, shown in Fig. 19, are overlain as white lines. Some prominent hotspots occur within the buried river channels, as in Figs. 21(e) and (f), while many other prominent hotspots occur outside of the channels, as in Fig. 21(c).

From Figs. 20(a)–(f), we do observe many more extended hotspots to the east and south-east in these images. These hotspots are consistently observable in several tracks and seem to occur along a bathymetric contour that is deeper but parallel to the 100-m water depth contour. These images should be compared with Fig. 17 where the returns line-up close to the 100-m water depth contour.

Several other prominent and discrete scattering events close to the 100-m water depth region were also imaged at site 1 as shown in Figs. 22(a) and (b). These images also show how a hotspot consistency chart can be used to resolve ambiguity. In the single ping image of Fig. 22(a) from track 73, a set of prominent and discrete scattering events are

charted to the north-west and another set of more elongated events to the south-east. Only one of these is the true set of scattered returns while the other is ambiguous. From Fig. 22(b), which shows the hotspot consistency chart for the track, the events from the south-east are consistently charted to the same area, leading to a strong reinforcement of these events in the hotspot chart. The events to the north-west are distributed over a wide area in the hotspot chart with little repeatability, showing that these events are the ambiguous ones. The true scattered returns therefore originate from the south-east in the region with water depth of between 80 to 100 m. Geophysical surveys at the location of these extended clutter returns did not find any subbottom features.²⁶

The measured clutter on all the tracks at site 1 is also statistically analyzed to compare the frequency of occurrence of repeatable clutter inside and outside of buried river channels. The analysis is similar to that done at site 2 and discussed in Sec. IV A 2. The area used in the analysis is the full region shown in Fig. 19 of $17.2 \times 13.3 \text{ km}^2$, which is the same as that in Figs. 21(a)–(f). The area occupied by the buried river channels within this box is approximately 9% of the total area of the box. The results are shown in Table II. The means of between 13% to 15% and standard deviations of between 7% to 14% show that there is no statistically significant difference between the frequency of occurrence of repeatable clutter charted within areas occupied by buried river channels and that charted outside of areas occupied by channels. This finding is similar to that at site 2.

V. SUMMARY AND DISCUSSION

At all the three sites investigated in the New Jersey Strataform area, a large number of clutter events were imaged throughout the experiment. Many often persisted along a track. These events were imaged in near-real-time over wide areas extending roughly 40 km in range from the bistatic sonar. Most of the clutter events were measured in areas where the bathymetry is locally level with slopes $< 0.5^\circ$. At site 2, these scattering events are sometimes as prominent and consistent as those from the calibrated targets deployed at a similar range from the bistatic sonar. Without prior knowledge, it would be impossible to distinguish the targets from these features.

Some of the high-level clutter events intersected buried river channels at sites 1 and 2. In general, the registration of clutter events with the buried river channels at a given location could not be maintained over variations in track orientation. Since none of the tracks were repeated, it is unknown whether or not the registration of the clutter for a given track with the channels is repeatable over time.

We consider two plausible causes for the variability in measured clutter with changes in track orientation. If the measured clutter is caused by scattering from geologic features of the subbottom such as buried river channels, its variability shows that the scattering process is highly dependent on the bistatic location and orientation of the source and receiver array relative to the geologic features. An analogy can be drawn with the “glints” observed on a spoon. Due to the curved surface of the spoon, the glints on a spoon migrate when the bistatic orientation of the observer or light

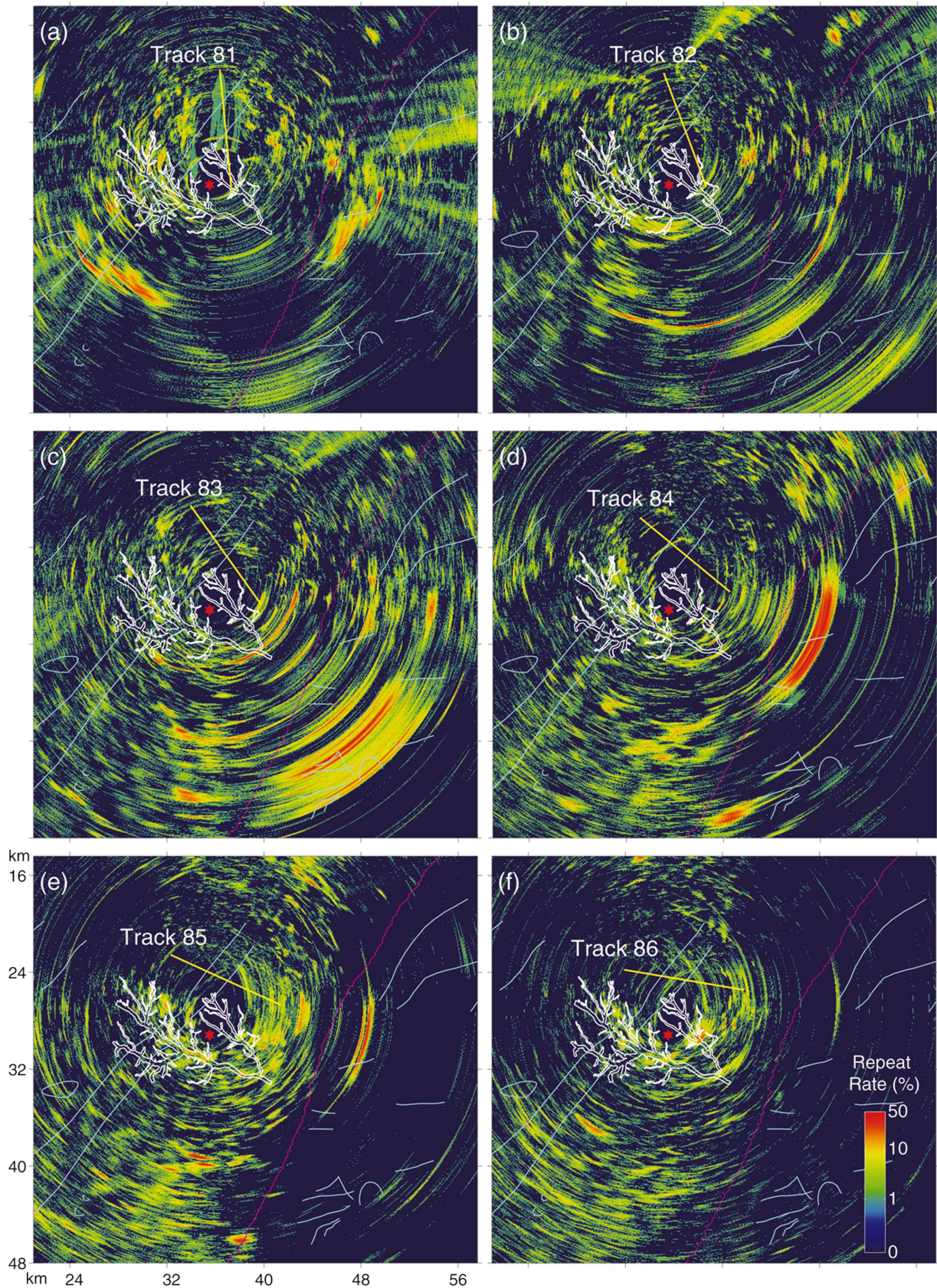


FIG. 20. Hotspot consistency chart for (a) 89 bistatic transmissions along track 81 with array heading 350° , (b) 78 bistatic transmissions along track 82 with array heading 154° , (c) 94 bistatic transmissions along track 83 with array heading 327° , (d) 73 bistatic transmissions along track 84 with array heading 122° , (e) 87 bistatic transmissions along track 85 with array heading 293° , and (f) 79 bistatic transmissions along track 86 with array heading 97° . The 100-m bathymetric contour is shown in pink. The most prominent and repeatable clutter, consistently observable in several of the tracks, occur in waters that is a little deeper than 100 m. The region where the buried river channels are shown is the only region where subbottom data have been collected and analyzed. There may be subbottom features in other regions shown in this figure, but the subbottom geologic data have either not been collected or not analyzed. [(a)–(d) Date: 4 May, mean source depth: 41 m, mean receiver depth: 35 m.]

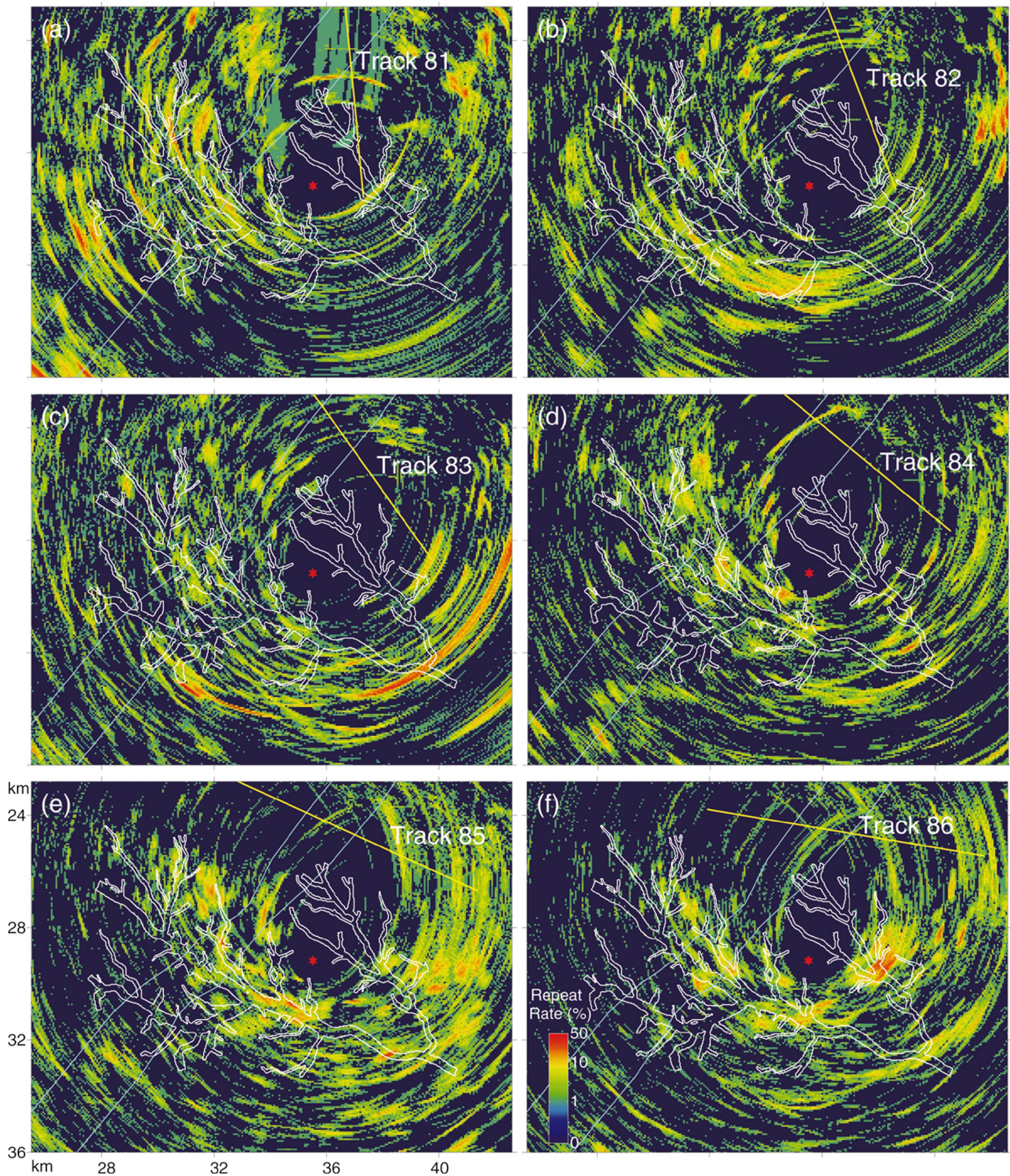


FIG. 21. (a)–(f) are similar to Fig. 20 but are zoomed in around the channels at site 1. The clutter registration with the channels changes with the orientation of the tracks which were done at different times of the day.

source changes.²⁹ The projected area of the imaged feature within the cross-range resolution of the measurement system changes with the bistatic orientation of the system relative to the feature. Bistatic aspects that lead to large projected areas would produce the strongest returns. In deep water, this is the case with deterministic reverberation from scarps on the Mid-Atlantic Ridge. The scattering strength of the scarp was found to be proportional to the projected area of the scarp

along the path from target to the bistatic source and receiver.^{8–10} Another possibility is that the clutter arises from objects moving around in the water column. This could account for the highly spatial and temporal variability observed in the measured clutter. Regardless of the origin of the clutter, returns in a waveguide also fluctuate as a result of propagation effects due to interference between the modes. The strength and consistency of the clutter measured at site 2 was

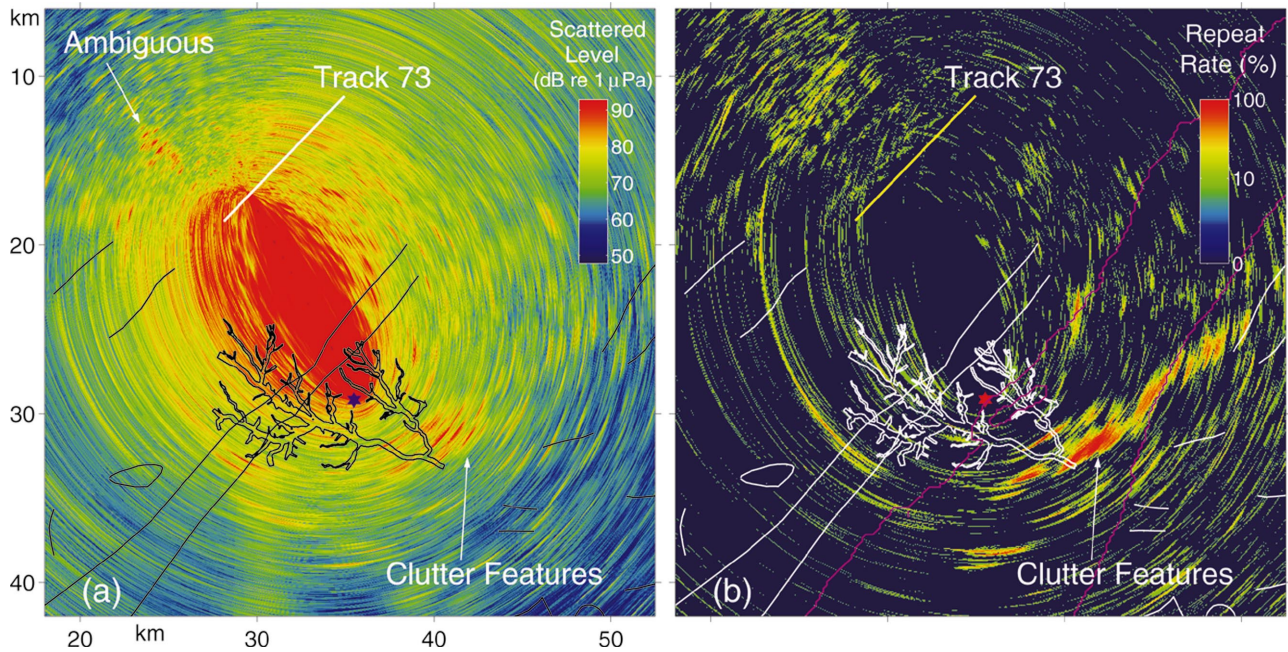


FIG. 22. (a) Single-ping bistatic wide-area image from track 73 at site 1. (b) Hotspot consistency chart for 12 bistatic transmissions along track 73 at site 1. In (a), a set of prominent and discrete scattering events has been charted to the north-west and another set of more elongated events have been charted to south-east. Only one of these is the true scattered return from the features while the other is ambiguous. Comparing (a) and (b) shows that the events to the south-west are the true returns. These returns do not coincide with any known geologic features, but are contained within the 80- to 100-m bathymetric contour (pink lines) of the STRATAFORM area at site 1. [Date: 3 May, (a) ping time: 16:28:35Z, transmission: 2 s duration LFM from 390 to 440 Hz. (a) and (b) Mean source depth: 41 m, mean receiver depth: 38 m, array heading: 220°.]

also found to vary with receiver depth and could be due to effects of the sound speed profile in the water column.

It should be noted that statistical analysis of the measured clutter at sites 1 and 2 in regions where the buried river channels have been mapped show that there is no significant difference between the rate of repeatable clutter per unit area charted within areas occupied by buried river channels and that charted outside of areas occupied by channels.

Seafloor features, such as the one iceberg scour and the large wall of the Hudson Canyon, are the only geologic features consistently imaged in the acoustic experiment. The one iceberg scour was detected at approximately 5-km range while the Canyon wall was imaged at about 30 km from the bistatic sonar at site 3. The returns from the Canyon wall are strong, but extended and not target-like in appearance.

Many of the prominent clutter events at all three sites occurred in the region along bathymetric contour between 80- and 120-m water depth. This is roughly the region where the warmer slope water mixes with the cooler shelf water in the New Jersey Strataform area as discussed in Appendix A.

One possibility for the source of unidentified clutter measured at sites 1–3 is scattering from fish.^{30–33} Fish are known to exist in tight schools about 5 m off the bottom in the STRATAFORM area with 80- to 120-m water depth.^{34,35} This is where the shelf break front exists which contains nutrients and is a source of food that attracts the fish schools. This possibility of fish schools causing clutter in long range sonar has been shown to be theoretically plausible^{30–33} and has also been observed experimentally.³⁶ When fish congregate together to form large and densely populated schools, they can scatter either coherently or incoherently, leading to returns that stand significantly above the diffuse reverbera-

tion background.^{32,33} This has been observed experimentally where such fish schools were found to be distributed in fan-shaped patterns.³⁶

Another possible origin for the unidentified clutter is

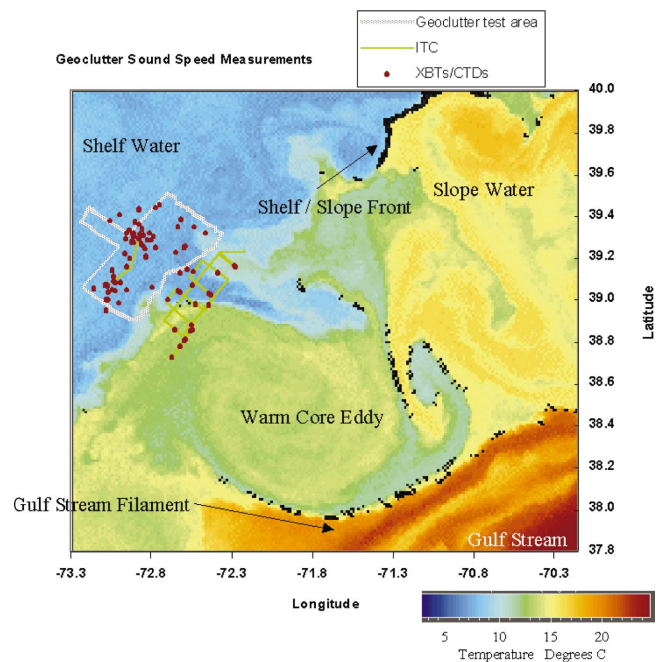


FIG. 23. Satellite image of the sea surface temperature off the New Jersey Coast on 30 April 2001. The oceanography in the region is governed by the interaction between the cooler shelf water (in blue), the warmer slope water (in yellow), and much warmer water from the gulf stream (in brown). Also visible is a warm core eddy (in green). Oceanographic data were collected using XBTs, CTDs, and an instrumented tow cable (TC) that allows almost continuous measurement of temperature along a tow cable.

TABLE II. Same as Table I but calculated for the tracks and the box with the full area in Fig. 19 at site 1. Buried river channels occupy 9.0% of the area of this box. There are a total of $M=21$ tracks at site 1 that transmitted bistatic LFM signals.

Track Index, n	Track name	N_n	$P_n(10\%)$ (%)	$P_n(20\%)$ (%)
1	1c	9	15.9	17.6
2	1x	14	22.9	31.7
3	2	11	11.4	9.5
4	3	16	9.4	7.4
5	4	15	3.3	3.3
6	5	15	26.5	41.6
7	7	15	13.4	9.0
8	8	14	15.5	15.3
9	11	12	8.9	7.2
10	12	13	12.4	15.5
11	71	11	8.8	5.1
12	72	15	5.3	3.1
13	73	12	6.0	4.8
14	74	12	7.4	6.5
15	75	18	12.9	9.9
16	81	89	8.5	3.1
17	82	78	7.2	0.0
18	83	94	7.5	7.0
19	84	73	7.7	0.0
20	85	87	23.6	45.5
21	86	79	30.3	43.1
Unweighted mean			12.6	13.6
Unweighted standard deviation			7.3	14.0
Weighted mean			13.5	15.4
Weighted standard deviation			8.5	10.3

scattering from gas entrapments (pockets) within the sediment that have significant acoustic impedance contrast with the neighboring sedimentation.³⁷ There is, however, much debate over whether these pockets of gases can exist in the East Coast STRATAFORM area. None of the geophysical surveys found any evidence of gas pockets in this region. Gas pockets are therefore less likely to be the source of the measured clutter in comparison to fish schools that have been proven to exist in abundance in this region.

The temporal and spatial variability of the clutter events, along with resonance characteristic of the clutter, provide important clues regarding the identity of the scatterers and the physical mechanisms associated with the scattering process. Scattering from geologic features is expected to be repeatable at a given location as a function of time. Clutter from fish schools, on the other hand, is expected to evolve and change with both space and time as the fish congregate and disperse. Detailed theoretical analysis of the scattering from fish schools, buried river channels, the R-reflector, and reverberation from the sea bottom on the New Jersey Continental Shelf will be presented in a separate paper.

In order to condense the information contained in data from large numbers of transmissions along a track-line of the receiver ship, a “hotspot” consistency chart was derived from the images of individual transmissions along each

track. The hotspot consistency chart displays the location and persistence of strong echo returns for transmissions along a track. In many instances, the hotspot charts provide significantly improved imaging of the clutter features. They are also useful for resolving the left–right ambiguity inherent in horizontal line array data when the bistatic range to the target is not much larger than the source–receiver separation. The hotspot charts also reduce charting errors due to waveguide dispersion since they combine information from a large number of transmissions.

The results of this experiment show that a low-frequency active sonar system can be used to remotely image the underwater environment over wide areas, at ranges spanning tens of kilometers, in continental shelf waters in near-real-time.

ACKNOWLEDGMENTS

The authors would like to thank Jefferey Simmen and Ellen Livingston who supported and sponsored the Acoustic Clutter Program at the U.S. Office of Naval Research. We would also like to thank NATO’s SACLANT Center for support of the experiment. We thank the scientists and crew of NRV *Alliance* and RV *Endeavor* for their effort in making the Acoustic Clutter Reconnaissance Experiment 2001 a suc-

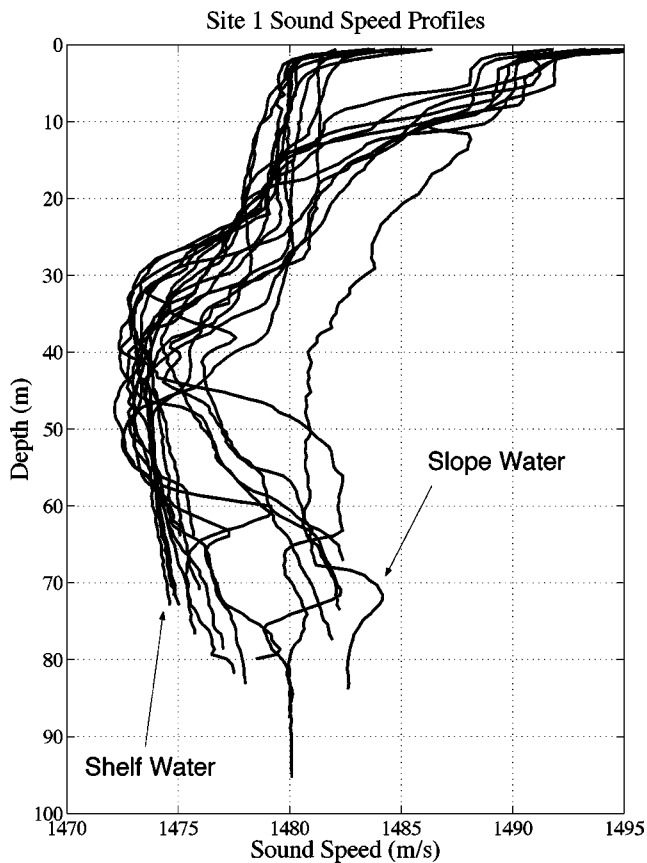


FIG. 24. Measured sound speed profiles of the water column at site 1.

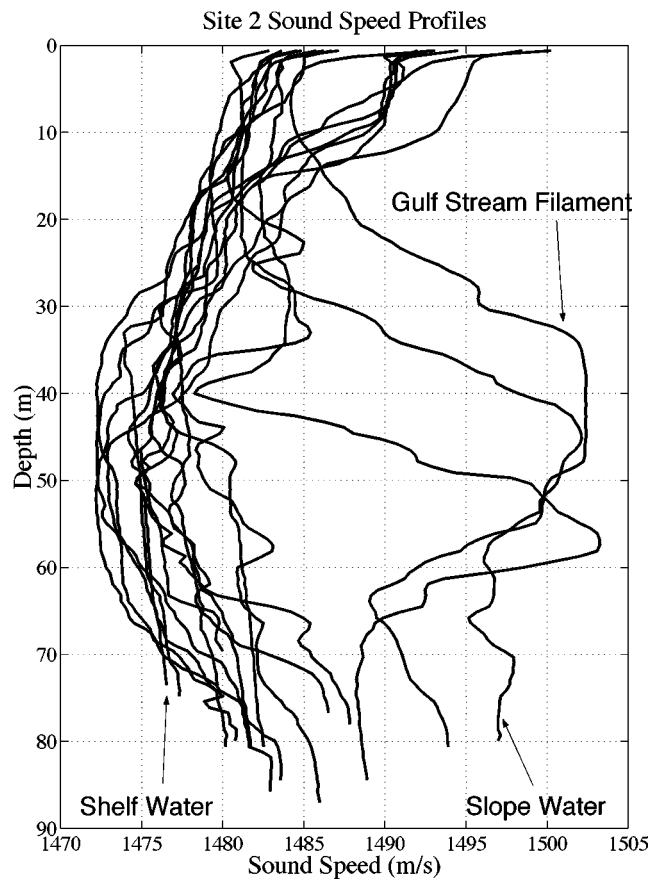


FIG. 25. Measured sound speed profiles of the water column at site 2.

cess. We would also like to thank Tommy Goldsberry for allowing us to use the MACE source and the engineering staff of the MACE Program for operation of the source during the experiment. We would like to acknowledge our colleagues at the University of Texas at Austin for providing us with geophysical support of the Program.

APPENDIX: STRATAFORM OCEANOGRAPHY

The East Coast STRATAFORM area has been fairly well characterized in general physical oceanography.³⁸⁻⁴⁰ The dominant water masses consist of seasonally dependent shelf water, the surface mixed layer, and continental slope water (Fig. 23). The boundary between shelf and slope water is known as the shelf break front, and consists of a strong salinity gradient with seasonally dependent temperature gradients. Its shape and location are variable at different times of the year. Its surface expression is usually near the 200- to 1000-m bathymetric contour, while its intersection with the bottom frequently manifests up onto the outer continental slope, as shallow as the 70-m contour. Thus the shelf break front and slope waters influence the lowermost part of the water column in the test area. The upper part during the experiment was dominated by surface warming, the mixed layer, and a thermocline down to the mid-water-column shelf water.⁴¹

The oceanographic measurements during the experiment consisted of XBTs from both the RV *Endeavor* and NRV *Alliance*, 27 Seabird CTD casts from the RV *Endeavor*, and five deployments of the Instrumented Tow Cable,⁴² a nearly

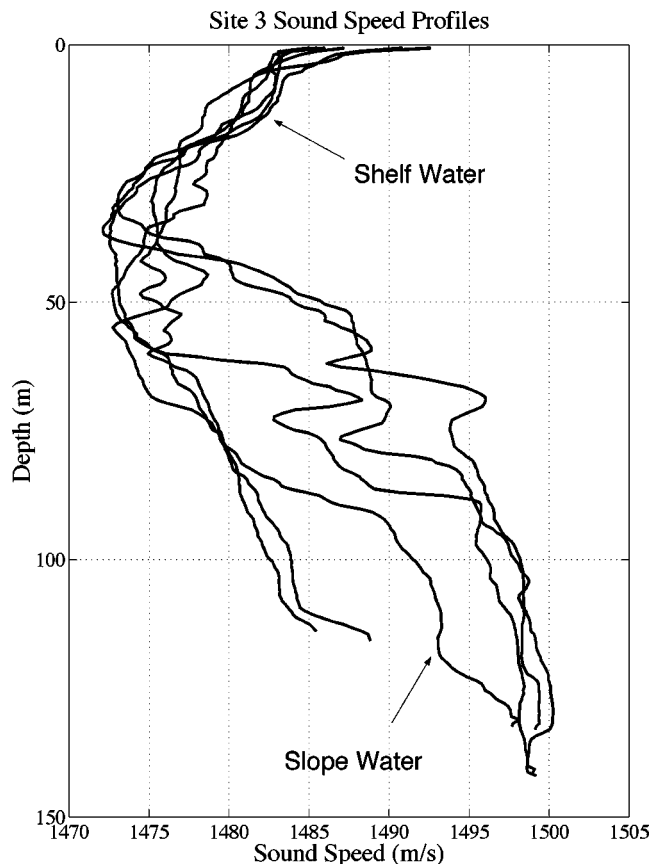


FIG. 26. Measured sound speed profiles of the water column at site 3.

continuous measurement device for temperature along a tow cable. In general, the measurements showed a sound speed profile with a moderately thin surface mixed layer (10 m or less), a thermocline of moderate gradient, to a temperature and sound speed minimum near 30- to 50-m depth (depending on the location) containing the “cold pool” of shelf water, below which was an increase in salinity and temperature associated with the shelf break front zone of mixing. The conditions have a tendency to channel the sound energy away from both the surface and bottom of the water column in the deeper portions of the area, and to allow better sound interaction with the bottom in areas shallower than 80 m.

Beyond this general physical oceanographic framework were two processes having a more variable effect on the sound speed structure. One was the existence of an unusually warm air mass over the area, causing near-surface warming continuously and gradually during the experiments, with a more pronounced effect in late afternoon. The second was the presence of a major (50 km diameter) slope eddy with an anticyclonal motion sitting just east of the shelf break front directly offshore of the STRATAFORM test area (Fig. 23). These warm eddies are known to have considerable influence on cross-frontal water mass movements and enhancements of chlorophyll in the shelf region.^{43–45} This feature had a fairly important and variable influence on the waters over the shelf up to the 85-m bathymetric contour during the experiments.

Figures 24–26 show representative sound speed profiles measured at sites 1–3, respectively. At site 1, Fig. 24 shows the effect of the near surface warming, a mixed layer less than 10 m thick, a fairly equal thermocline and salinity front, with the cold pool of shelf water on the order of 10 to 20 m thick centered on 40-m water depth. The deeper water profiles show a strong influence of the shelf slope front, while the probes from 70 m and shallower show no influence at all from the slope waters, with the cold pool extending to the bottom.

At site 2, Fig. 25, three of the sound speed profiles show the presence of entrained gulf stream waters in the mid water column. These three profiles were the furthest east collected during the experiment. The warm waters from the gulf stream filament produced a high-speed sound layer in the water column for some areas deeper than the 80-m bathymetric contour. Outside the region of influence of this filament, however, there was the more normal behavior of the cold pool extending to the bottom west of the 70-m contour and the influence of the shelf break front deeper than 50 to 60 m in the water column. The thermocline moved up or down by ± 5 m, while the lower to mid part of the water column was influenced by various filaments of warm water, creating a less stable sound speed profile than at site 1.

Site 3 was further towards the shelf edge, and measurements were sampled deeper than at the other two sites. In Fig. 26, the thermocline moved up and down on the order of ± 5 m. There were cooler waters near the bottom for profiles taken shoreward of the mooring site, and a dominant presence of slope water in the near-bottom profiles further offshore. In these profiles slope waters themselves appear deeper than 95 m in the water column. This causes a fairly sharp upward refracting profile below the sound channel

axis. This caused less bottom interaction at site 3. Finally, in several profiles we observe the influence of fingers or filaments of warmer waters in the mid-water column.

The oceanographic measurements at the East Coast STRATAFORM show the interaction between the various water masses that influence the outer shelf water properties. This highlights the need for making high resolution temperature measurements, especially for understanding their effects on acoustic propagation in this region.

- ¹N. C. Makris, Editor “Geological Clutter Acoustics Experiment,” Office of Naval Research Initial Report, Geoclutter Program, 27 April–5 May 2001.
- ²N. C. Makris and P. Ratilal, “A unified model for reverberation and submerged object scattering in a stratified ocean waveguide,” *J. Acoust. Soc. Am.* **109**, 909–941 (2001).
- ³P. Ratilal, N. C. Makris, and Y. Lai, “Validity of the sonar equation and Babinet’s principle for scattering in a stratified medium,” *J. Acoust. Soc. Am.* **112**, 1797–1816 (2002).
- ⁴F. Ingenito, “Scattering from an object in a stratified medium,” *J. Acoust. Soc. Am.* **82**, 2051–2059 (1987).
- ⁵N. C. Makris, “A spectral approach to 3-D object scattering in layered media applied to scattering from submerged spheres,” *J. Acoust. Soc. Am.* **104**, 2105–2113 (1998).
- ⁶N. C. Makris, F. Ingenito, and W. A. Kuperman, “Detection of a submerged object insonified by surface noise in an ocean waveguide,” *J. Acoust. Soc. Am.* **96**, 1703–1724 (1994).
- ⁷F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, New York, 1994).
- ⁸N. C. Makris, L. Z. Avelino, and R. Menis, “Deterministic reverberation from ocean ridges,” *J. Acoust. Soc. Am.* **97**, 3547–3574 (1995).
- ⁹N. C. Makris and C. S. Chia, “The bi-azimuthal scattering distribution of an abyssal hill,” *J. Acoust. Soc. Am.* **106**, 2491–2512 (1999).
- ¹⁰C. S. Chia, N. C. Makris, and L. T. Fialkowski, “A comparison of bistatic scattering from two geologically distinct abyssal hills,” *J. Acoust. Soc. Am.* **108**, 2053–2070 (2000).
- ¹¹J. R. Preston and T. Akal, “Analysis of backscattering data in the Tyrrhenian Sea,” *J. Acoust. Soc. Am.* **87**, 119–134 (1990).
- ¹²M. D. Max, N. Portunato, and G. Murdoch, “Sub-seafloor buried reflectors imaged by low frequency active sonar,” SACLANTCEN memorandum, SM-306, SACLANT Centre, La Spezia, Italy (1996).
- ¹³C. A. Nittrouer, “STRATAFORM: overview of its design and synthesis of its results,” *Mar. Geol.* **154**(1–4), 3–12 (1999).
- ¹⁴C. S. Duncan, J. A. Goff, J. A. Austin, and C. S. Fulthorpe, “Tracking the last sea-level cycle: seafloor morphology and shallow stratigraphy of the latest Quaternary New Jersey middle continental shelf,” *Mar. Geol.* **170**, 395–421 (2000).
- ¹⁵J. A. Goff, D. Swift, C. S. Duncan, L. A. Mayer, and J. Hughes-Clarke, “High-resolution swath sonar investigation of sand ridge, dune and ribbon morphology in the offshore environment of the New Jersey margin,” *Mar. Geol.* **161**, 307–337 (1999).
- ¹⁶C. S. Duncan and J. A. Goff, “Relict iceberg keel marks on the New Jersey outer shelf, southern Hudson apron,” *Geology* **29**, 411–414 (2001).
- ¹⁷R. B. Evans and W. M. Carey, “Reproducibility of low-frequency shallow-water acoustic experimental data,” *J. Acoust. Soc. Am.* **99**, 2523–2529 (1996).
- ¹⁸K. M. Becker, G. V. Frisk, and S. D. Rajan, “Geoacoustic inversion results from the modal mapping experiment,” *J. Acoust. Soc. Am.* **107**, 2775 (2000).
- ¹⁹A. Turgut and S. Wolf, “Matched-field inversion of seabed geoacoustic properties complemented by chirp sonar surveys,” *J. Acoust. Soc. Am.* **110**, 2661 (2001).
- ²⁰C. I. Malme, P. Jameson, P. McElroy, D. Stracher, G. Thomas, and D. Zwillinger, “Development of a High Target Strength Passive Acoustic Reflector for Low Frequency Sonar Applications,” BBN Report No. 7943. BBN Technologies, New London, CT.
- ²¹N. C. Makris and J. M. Berkson, “Long-range backscatter from the Mid-Atlantic Ridge,” *J. Acoust. Soc. Am.* **95**, 1865–1881 (1994).
- ²²N. C. Makris, “Imaging ocean-basin reverberation via inversion,” *J. Acoust. Soc. Am.* **94**, 983–993 (1993).
- ²³G. Bar-Yehoshua, “Quantifying the effect of dispersion in continental shelf sound propagation,” Master’s thesis, MIT, 2002.

- ²⁴J. W. Goodman, *Statistical Optics* (Wiley, New York, 1985).
- ²⁵N. C. Makris, "The effect of saturated transmission scintillation on ocean acoustic intensity measurements," *J. Acoust. Soc. Am.* **100**, 769–783 (1996).
- ²⁶J. Austin, J. Goff, S. Gulick, C. Fulthorpe, S. Nordfjord, M. Wiederspahn, S. Sastrup, S. Schock, J. Wulf, K. Gjerding, L. Mayer, and C. Sommerfield, "Assessing the "Geo,"" in *GEOCLUTTER: New chirp sonar, sampling, and compressional wave velocity results from the New Jersey shelf*," *Eos. Trans. AGU.* **82**(47), Fall Meet. Suppl. Abstract OS42A-0456 (2001).
- ²⁷C. S. Fulthorpe, J. A. Goff, J. A. Austin, S. P. Gulick, and S. Nordfjord, "Late Quaternary Incisions and Related Shallow Subsurface Stratigraphy on the New Jersey Mid-Outer Shelf: Preliminary Results from Ultra-High Resolution Chirp Sonar Images—Part II," *EOS Trans. AGU.* **83**, Fall Meet. Suppl., Abstract OS71C-0300 (2001).
- ²⁸J. A. Goff and S. Nordfjord, "Interpolation of fluvial morphology using channel-oriented coordinate transformation: A case study from the New Jersey shelf," submitted to *Mathematical Geology*.
- ²⁹N. C. Makris, P. Ratilal, Y. Lai, and E. K. Scheer, "The Geoclutter experiment: Remotely imaging sub-bottom geomorphology in shallow water," *J. Acoust. Soc. Am.* **110**, 2742 (2001).
- ³⁰R. C. Gauss, R. W. Nero, and D. Wurmser, "A Lloyd-mirror model to estimate the scattering strength of fish near rough ocean boundaries," *J. Acoust. Soc. Am.* **104**, 1820(A) (1998).
- ³¹R. Gauss, R. F. Gragg, R. W. Nero, D. Wurmser, and J. M. Fialkowski, "Broadband models for predicting bistatic bottom, surface, and volume scattering strengths," *NRL Report, NRL/FR/7100-02-10,042* (2002).
- ³²P. Ratilal, "Remote sensing of submerged objects and geomorphology in continental shelf waters with acoustic waveguide scattering," Ph.D Thesis, MIT, 2002.
- ³³P. Ratilal and N. C. Makris, "Coherent versus diffuse surface and volume reverberation in an ocean wave guide: Reverberation rings, modal decoupling, and possible fish scattering in Geoclutter 2001," *J. Acoust. Soc. Am.* **112**, 2280 (2002).
- ³⁴R. W. Nero, "Fish scattering during Boundary Characterization 2001," in *Proceedings of Geoclutter and Boundary Characterization 2001: Acoustic Interaction with the Seabed*, Halifax, October 2001, edited by P. C. Hines, N. C. Makris, and C. W. Holland, published by the Defence Research Establishment Atlantic, DREA TM-2001-185 (2001), pp. 39–42.
- ³⁵R. W. Nero, C. H. Thompson, and R. H. Love, "Fish schools as potential clutter and false targets: Observations on the New Jersey shelf," *J. Acoust. Soc. Am.* **112**, 2280–2281 (2002).
- ³⁶D. E. Weston and J. Revie, "A 5-day long-range sonar record of an extensive concentration of fish," *J. Acoust. Soc. Am.* **86**, 1608–1611 (1989).
- ³⁷C. Holland, "Self-Consistent Measurements of Seabed Reflection and Scattering in the STRATAFORM Area," in *Proceedings of Geoclutter and Boundary Characterization 2001: Acoustic Interaction with the Seabed*, Halifax, October 2001, edited by P. C. Hines, N. C. Makris, and C. W. Holland, published by the Defence Research Establishment Atlantic, DREA TM-2001-185 (2001), pp. 12–20.
- ³⁸G. Gawarkiewicz, T. G. Ferdelman, T. M. Church, and G. W. Luther III, "Shelf break frontal structure on the continental shelf north of Cape Hatteras," *Cont. Shelf Res.* **16**, 1751–1773 (1996).
- ³⁹R. W. Houghton, F. Aikman III, and H. W. Ou, "Shelf-slope frontal structure and cross-shelf exchange at the New England shelf break," *Cont. Shelf Res.* **8**, 687–710 (1988).
- ⁴⁰C. A. Linder and G. Gawarkiewicz, "A climatology of the shelfbreak front in the Middle Atlantic Bight," *J. Geophys. Res.* **103**, 18405–18424 (1998).
- ⁴¹R. W. Houghton, R. Schlitz, R. C. Beardsley, B. Bufman, and J. L. Chamberlain, "The Middle Atlantic Bight cold pool: evolution of the temperature structure during summer 1979," *J. Phys. Oceanogr.* **12**, 1019–1029 (1982).
- ⁴²A. Ruffa, J. Bard, and M. Sundvik, "Demonstration test of LBVDS instrumented tow cable in waters South of New England—Initial Results," in *Oceans 2000 Conference Proceedings*, Providence, RI, October 2000.
- ⁴³J. A. Barth, D. Bogucki, S. D. Pierce, and P. M. Kosro, "Secondary circulation associated with a shelfbreak front," *Geophys. Res. Lett.* **25**, 2761–2764 (1998).
- ⁴⁴G. Gawarkiewicz, F. Bahr, R. C. Beardsley, and K. H. Brink, "Interaction of a slope eddy with the shelfbreak front in the Middle Atlantic Bight," *J. Phys. Oceanogr.* **31**, 2783–2796 (2001).
- ⁴⁵J. P. Ryan, J. A. Yoder, and D. W. Townsend, "Influence of a Gulf Stream warm-core ring on water mass and chlorophyll distributions along the southern flank of Georges Bank," *Deep-Sea Res., Part II* **48**, 159–178 (2000).

Time-reversal imaging for classification of submerged elastic targets via Gibbs sampling and the Relevance Vector Machine

Nilanjan Dasgupta and Lawrence Carin^{a)}

Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina 27708-0291

(Received 19 April 2004; revised 15 December 2004; accepted 8 January 2005)

Time-reversal imaging (TRI) is analogous to matched-field processing, although TRI is typically very wideband and is appropriate for subsequent target classification (in addition to localization). Time-reversal techniques, as applied to acoustic target classification, are highly sensitive to channel mismatch. Hence, it is crucial to estimate the channel parameters before time-reversal imaging is performed. The channel-parameter statistics are estimated here by applying a geoacoustic inversion technique based on Gibbs sampling. The maximum *a posteriori* (MAP) estimate of the channel parameters are then used to perform time-reversal imaging. Time-reversal implementation requires a fast forward model, implemented here by a normal-mode framework. In addition to imaging, extraction of features from the time-reversed images is explored, with these applied to subsequent target classification. The classification of time-reversed signatures is performed by the relevance vector machine (RVM). The efficacy of the technique is analyzed on simulated in-channel data generated by a free-field finite element method (FEM) code, in conjunction with a channel propagation model, wherein the final classification performance is demonstrated to be relatively insensitive to the associated channel parameters. The underlying theory of Gibbs sampling and TRI are presented along with the feature extraction and target classification via the RVM. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1862592]

PACS numbers: 43.30.Pc, 43.60.Pt, 43.60.Lq, 43.60.Uv [DRD]

Pages: 1999–2011

I. INTRODUCTION

The problem of imaging and classifying a target situated in a shallow-water (sound) channel is complicated by multipath, due to multiple reflections within the sound channel. Such multipath is dependent on parameters such as the target and receiver location, the channel depth, the sound-speed profile within the channel, and the geoacoustic parameters of the channel bottom.¹ The impact of the channel on the measured target signature may be significant. Therefore, it is often important to account for channel effects on the target signature² when performing classification. Our objective is to remove the channel effects from the in-channel target response such that a channel-independent classifier can be designed for classification of submerged in-channel targets.

Pioneering work on classification of underwater targets was detailed by Gorman and Sejnowski,³ where they used spectral features to distinguish cylindrical targets based on a neural-network classifier. Goo *et al.*⁴ have developed the “G-Transform” which, in conjugation with a neural-network-based classifier, worked successfully to detect and identify targets in shallow water. These techniques try to learn the target response in the sound channel; hence, the classifier performance is severely degraded when the anticipated and actual parameters are mismatched. Consequently, an important issue in sensing underwater targets involves removing the effects of the channel. This is motivated by the flexibility of using a single classifier, trained on free-field target signatures, for classifying the scattered responses from the same target(s) situated in different acoustic environments—by

contrast, it is prohibitive to require a different classifier for each channel environment. The deterministic deconvolution method for extracting the scattered signature is very sensitive to channel Green’s function mismatch. Other researchers have employed statistical techniques, like higher order statistics⁵ and minimum entropy deconvolution (MED).⁶ For the case of a shallow water channel with long observation distances (for which many modes are excited), one requires a large array of receivers for the channel Green’s function to satisfy the requirements of these methods. In addition to work on channel deconvolution, Runkle *et al.*⁷ have demonstrated that classification accuracy may be greatly improved by fusing scattering data from multiple orientations around a target, via a hidden Markov model (HMM).

In the work reported here time-reversal imaging is applied to simulated data, from which features are extracted; these features are processed subsequently to make a classification decision. Time reversal is manifested by taking the time-domain scattered fields received on an array, reversing them in time, and then propagating them back into the domain. One must distinguish active time-reversal from time-reversal imaging (TRI). In active time-reversal⁸ the signals are *physically* radiated into the domain via a transducer, and the spatio-temporal focusing at the source may be used for covert communication and/or to direct a large amount of energy at the source. In TRI the time-reversed signals are radiated into the domain *computationally*, yielding a space-time image, with temporal and spatial focusing at the scatterer. The main distinction between TRI and conventional matched-field processing⁹ is that TRI is typically wideband^{10,11} and implemented in the time domain (i.e., the frequencies are processed coherently). Jackson and

^{a)}Electronic mail: lcarin@ee.duke.edu

Dowling¹² have established a theoretical understanding of the time-reversal technique applied to underwater imaging, considering a simulated medium. The time-reversal image yields a four-dimensional signal (three spatial dimensions, and time), and the time dependence of each scattering center is also recovered approximately. The spatio-temporal focusing achieved by time-reversal imaging is strongly dependent on the similarity between the true and simulated acoustic channel (the former is characteristic of the scattering data in the measurements, and the latter is employed computationally in the TRI procedure).⁸

There have been several manifestations of time-reversal imaging in the literature. For example, Fink and colleagues¹³ have performed an eigen-decomposition of the time-reversal operator, and each eigenmode may be linked to a specific target or subtarget. These eigenmodes may be radiated actively or numerically into the domain, to achieve selective focusing. These eigenmodes are closely related to the modes in a MUSIC algorithm¹⁴ and authors have also considered¹⁵ time-reversal imaging in a modified MUSIC framework. Carin *et al.*^{1,16} have employed direct space-time time-reversal imaging to measured underwater scattering data, including estimation of the unknown channel parameters via a genetic-algorithm optimization applied to the time-reversal image.

As discussed by Carin *et al.*,^{1,16} the quality of the time-reversal imagery is dependent on the accuracy with which the channel Green's function is estimated. This is closely related to previous research in matched-field processing. Dosso^{17,18} has demonstrated that the quality of focusing in acoustic matched-field processing is a strong function of the mismatch between the numerical and actual channel parameters. This dependence of quality on channel-parameter mismatch has been exploited to invert for the channel parameters using a Gibbs-sampling approach (GS).¹⁹ Gibbs sampling is used in a Bayesian framework to generate the *a posteriori* distribution of the channel parameters, based on the observed data and a statistical prior distribution over the channel parameters (if available). In this paper time-reversal imaging²⁰ of underwater targets is performed based on channel-parameter statistics estimated via GS.

After performing Gibbs sampling to estimate channel parameters, performing time-reversal imaging, and then extracting features from the time-reversal imagery, the features are utilized to make a classification decision. A nonlinear classifier called the relevance vector machine (RVM) is employed to classify targets from (non-targets) in feature space. The RVM was first proposed by Tipping²¹ as a Bayesian generalization of the support vector machine (SVM).²² The RVM is an algorithm for finding the best Gaussian prior distribution for a logistic regression (LR)²³ classifier, by using evidence maximization²¹ (also called the type-II maximum likelihood). The objective is to estimate a robust and generalized classification in the feature space.

This paper therefore represents an integration of several state-of-the-art tools in imaging, statistics, and classifier design. The space-time image is manifested via TRI, and the features from the TRI image are processed to yield a classification decision, based on the RVM. Finally, the quality of

the TRI algorithm and associated features employed within the RVM is dependent upon the accuracy with which the channel Green's function is estimated. Rather than simply performing a maximum-likelihood estimate of the channel parameters (point estimate), Gibbs sampling is used to estimate the full posterior distribution of the model parameters, the results from which are integrated into the TRI algorithm.

The remainder of the text is organized as follows. In Sec. II we briefly address the theory of time-reversal imaging, and in Sec. III Gibbs sampling is explained in the context of channel-parameter estimation. Feature extraction and classifier design based on the RVM are discussed in Sec. IV. Target classification performance is analyzed in Sec. V, followed by conclusions in Sec. VI.

II. TIME-REVERSAL THEORY

The phenomenon of time-reversal has been studied by several scientists^{8,11,24-26} in actual channel environments and for simulated acoustic medium. Previous researchers^{12,26} have presented the physics of time-reversal; hence the phenomenon is briefly described here, concentrating on the specifics of our problem. For an acoustic waveguide, scattering center localization can be achieved by spanning the channel by a vertical array of transceivers (say, L_r receivers, at locations \mathbf{r}_{rl} , $l \in \{1, \dots, L_r\}$) with maximum interelement distance of $\lambda/2$ with λ being the minimum wavelength of the transmitted signal. Let \mathbf{r}' and \mathbf{r}_{sk} represent respectively the source and k th scattering center location. For a transmitted signal $S(\omega)$ from the source, the backscattered signal received by the l th receiver located at \mathbf{r}_{rl} is given by

$$R(\omega, \mathbf{r}_{rl}) = \sum_k S(\omega) \Gamma_k(\omega) G(\omega, \mathbf{r}_{sk}, \mathbf{r}') G(\omega, \mathbf{r}_{rl}, \mathbf{r}_{sk}), \quad (1)$$

where Γ_k represents the reflectivity of the k th scattering center, $G(\omega, \mathbf{r}, \mathbf{r}')$ represents wave propagation from location \mathbf{r} to \mathbf{r}' , and the summation is over all scattering center locations. The target scattering is represented in terms of diffraction from isolated scattering centers; a more-general analysis has been performed by Carin *et al.*¹ Launching the time-reversed signal into the waveguide is equivalent to phase conjugation in the frequency domain. The phase-conjugated signal observed after propagating from \mathbf{r}_{rl} to any observation point \mathbf{r} is expressed as

$$S_{TR}(\omega, \mathbf{r}) = R^*(\omega, \mathbf{r}_{rl}) \hat{G}(\omega, \mathbf{r}, \mathbf{r}_{rl}), \quad (2)$$

where $\hat{G}(\omega, \mathbf{r}, \mathbf{r}')$ represents the numerically computed Green's function between \mathbf{r} and \mathbf{r}' . Assuming perfect knowledge about the channel, one can express $\hat{G}(\omega, \mathbf{r}, \mathbf{r}') = G(\omega, \mathbf{r}, \mathbf{r}')$, and due to reciprocity $G(\omega, \mathbf{r}_{rl}, \mathbf{r}_{sk}) = G(\omega, \mathbf{r}_{sk}, \mathbf{r}_{rl})$. Therefore constructive interference (focusing) is obtained at $\mathbf{r} = \mathbf{r}_{sk}$, with the wrong time dependence due to the effects of the initial propagation through the channel, from the source to the respective scattering center. To eliminate the effects of this propagation, one needs to run the forward model again to approximate $G(\omega, \mathbf{r}_{sk}, \mathbf{r}')$, which finally yields the time-dependent reflection coefficient, con-

volved with the incident pulse, at the point of the scattering center as

$$S_{TRM}(\omega, \mathbf{r}) = R^*(\omega, \mathbf{r}_{r_l}) \hat{G}(\omega, \mathbf{r}, \mathbf{r}_{r_l}) \hat{G}(\omega, \mathbf{r}', \mathbf{r}). \quad (3)$$

Repeating this same process for all the receivers and summing yields

$$S_{TRM}^{array}(\omega, \mathbf{r}) = \sum_{l=1}^{L_r} R^*(\omega, \mathbf{r}_{r_l}) \hat{G}(\omega, \mathbf{r}, \mathbf{r}_{r_l}) \hat{G}(\omega, \mathbf{r}', \mathbf{r}). \quad (4)$$

Performing an inverse Fourier transform of (4) yields a four-dimensional (space-time) image. This localizes the scattering centers in three-dimensional space while also approximately recovering the time dependence of each scattering center.

The above theoretical discussion assumes that the Green's function of the channel is known exactly. Typically, when performing TRI the Green's function is not known precisely *a priori*. Hence, one must perform geoacoustic inversion in order to estimate the channel parameters. In this paper a Gibbs sampling¹⁹ approach is employed to estimate these parameters.

III. CHANNEL INVERSION BY GIBBS SAMPLING

Dosso¹⁷ has developed a new Bayesian approach for estimating the geoacoustic parameters of a channel. Given prior knowledge of the model parameters, the proposed method estimates the posterior distribution based on Gibbs sampling.¹⁹ Since GS is explained in detail by Dosso,¹⁷ a brief summary is provided here in the context of our problem.

Let the vectors \mathbf{m} and \mathbf{d}^{obs} represent respectively the channel model parameters and observed scattering data at the array of receivers. Our objective is to estimate the underlying channel parameter vector \mathbf{m} given \mathbf{d}^{obs} . The *a posteriori* probability of the channel parameters $p(\mathbf{m}|\mathbf{d}^{\text{obs}})$ is of the form

$$p(\mathbf{m}|\mathbf{d}^{\text{obs}}) = \exp[-E(\mathbf{m})] p(\mathbf{m}) / \int_{\mathbf{m}'} \exp[-E(\mathbf{m}')] p(\mathbf{m}') d\mathbf{m}', \quad (5)$$

where $p(\mathbf{m})$ represents the prior distribution on the channel model parameters, and $E(\mathbf{m})$ is an energy function defined by Dosso¹⁷ as

$$E(\mathbf{m}) = \sum_{f=1}^F B_f(\mathbf{m}) |\mathbf{d}_f^{\text{obs}}|^2 / \eta, \quad (6)$$

where $B_f(\mathbf{m})$ represents the normalized Bartlett mismatch function $B_f(\mathbf{m}) = 1 - |\boldsymbol{\rho}_f^H(\mathbf{m}) \mathbf{d}_f^{\text{obs}}| / |\boldsymbol{\rho}_f(\mathbf{m})|^2 |\mathbf{d}_f^{\text{obs}}|^2$, with H being the complex conjugate transpose operator. The vector $\mathbf{d}_f^{\text{obs}}$ represents the backscattered signal at the receiver array for frequency f , and the parameter η represents the noise variance (frequency invariant, in our case). The vector $\boldsymbol{\rho}_f(\mathbf{m})$ is the simulated channel response as viewed by the array of receivers, for model parameters \mathbf{m} , due to a point source located at the position of the target.

Our objective is to approximate the integral in the denominator of (5) in terms of a sum, using a finite set of "important" examples of \mathbf{m} . This is executed by using a

TABLE I. The channel parameters estimated via the Gibbs sampling approach.

	True value	Estimated value	Search range
Channel depth (m)	6	6.01	(5.0–10.0)
SS in water (m/s)	1520	1523.2	(1470–1550)
SS in bottom (m/s)	1800	1807.8	(1750–1850)
Density of bottom (g/cm ³)	1.5	1.63	(1.0–2.0)
Attenuation of bottom (dB/kmHz)	0.28	0.295	(0.25–0.35)

variation of the Metropolis^{27,28} algorithm (MA). Assume that \mathbf{m}_i are the model parameters considered for iteration i , with associated energy $E(\mathbf{m}_i)$. Model parameters \mathbf{m}_{i+1} are deemed important if $\exp\{-[E(\mathbf{m}_{i+1}) - E(\mathbf{m}_i)]/T\} > \delta$, for temperature T , where δ is a random variable drawn from a uniform distribution over zero to one. Therefore, if $E(\mathbf{m}_{i+1}) < E(\mathbf{m}_i)$, then model parameter vector \mathbf{m}_{i+1} is always deemed important, and if $E(\mathbf{m}_{i+1}) > E(\mathbf{m}_i)$, the probability that \mathbf{m}_{i+1} is important is $\exp\{-[E(\mathbf{m}_{i+1}) - E(\mathbf{m}_i)]/T\}$. Therefore, with decreasing T , the probability of accepting parameters \mathbf{m}_{i+1} that increase the energy becomes very small.

In our problem, the channel model parameters \mathbf{m} are drawn randomly from the channel parameter search space using a uniform probability distribution [$p(\mathbf{m})$ is uniform]. The search space is initialized based on the prior knowledge of the channel environment (see Table I for the search space used in our problem). For each sample \mathbf{m}_i drawn uniformly from the model search space, the impulse response of the corresponding channel $\boldsymbol{\rho}_f(\mathbf{m})$ is computed numerically in a normal mode framework.²⁹ Given the impulse response $\boldsymbol{\rho}_f(\mathbf{m})$, the mismatch function $B_f(\mathbf{m})$, and the corresponding energy function, $E(\mathbf{m})$ is computed as shown in Eq. (6) (η is assumed known). The sample \mathbf{m}_i is accepted (or deemed important) if $E(\mathbf{m}_i) < E(\mathbf{m}_{i-1})$ or $\exp\{-[E(\mathbf{m}_i) - E(\mathbf{m}_{i-1})]/T\} > \delta$ with $p(\delta) = 1$, $\delta \in [0, 1]$. Note that the algorithm is started at $T > 1$ (T is initialized to 10). In subsequent iterations the temperature is lowered ($T^{\text{new}} = T^{\text{old}}/2$) until the temperature reaches unity and is kept constant thereafter. This process of accepting "important" samples is repeated until the resulting channel parameter posterior distribution converges. The "cooling" procedure is employed to achieve rapid convergence of the channel-parameter posterior distribution. In addition, the channel model search space is adaptively modified after every 20 new samples are accepted (deemed important). The modified range spans between $\boldsymbol{\mu}_m - 2\boldsymbol{\sigma}_m$ to $\boldsymbol{\mu}_m + 2\boldsymbol{\sigma}_m$, where $\boldsymbol{\mu}_m$ and $\boldsymbol{\sigma}_m$ represents the mean and standard deviation of the model parameters evaluated from the last 50 "important" samples accepted by MA.

IV. FEATURE EXTRACTION AND CLASSIFIER DESIGN

In the previous section a procedure for estimating Eq. (5) (the likelihood of the channel parameters given the observed data) is summarized. Time-reversal imaging can now be performed based on these estimated channel statistics. It is observed that, for relatively high signal-to-noise ratio (SNR),

Eq. (5) is typically strongly peaked about the true value for the channel parameters, and therefore the MAP channel estimator is employed in forming a time-reversal image [rather than integrating across the likelihood in Eq. (5), as in a Bayesian setting]. Details on the dependence of channel-parameter estimation on SNR are described when presenting example results.

The data under consideration were generated as follows. An FEM^{30,31} model was used to compute the multistatic scattered fields from five elastic targets in free field. The in-channel signatures were then computed by coupling these free-field solutions to a numerical forward model for the channel. The numerical details of the FEM model have been discussed previously.^{32,33}

The underwater targets possessed bisectonal symmetry, and therefore data were computed for 180° of variation in the target-sensor orientation, in 1° increments. A vertical receiver array was employed, and the source is placed at the array center. The observed array data for a given target-sensor orientation is time-reversed to generate a three-dimensional image in range, depth, and time (no azimuthal resolution). Target classification techniques applied to the time-reversed images are discussed in the following subsections.

A. Features

It has been demonstrated^{34,35} that, for the TRI procedure discussed in Sec. II, the image at $t=0$ corresponds to focusing at the target scattering centers. Feature are therefore extracted from the time-reversal image at time $t=0$, now corresponding to a range-depth image. Finally, to simplify things further, the depth is fixed, and the time-reversed one-dimensional signal is viewed as a function of range, through the nominal target center (which may be estimated from the range-depth two-dimensional image, prior to feature extraction). In the work presented here this process is automated, based on the energy distribution in the time-reversal image. Assume that $p_j(i,n)$ represents the one-dimensional time-dependent time-reversed data for target j at orientation n , where i denotes the time index (the time-dependent signal is the range-dependent waveform divided by the sound speed at the target depth). Features based on the spectral (Fourier) properties of $p_j(i,n)$ are considered in these examples for training the classifier. In this case the Fourier transform of $p_j(i,n)$, operated along dimension i , is uniformly divided into a set of seven frequency subbands, and the features are the energy within each band. The seven-dimensional feature vector \mathbf{v}_n corresponds to the orientation n of the target with respect to the vertical array. For the sake of simplicity the subscript j that represents the underlying target is ignored in the rest of the paper. It should be emphasized that only a subset of the information available in the time-reversal data is exploited here, since the image is a four-dimensional space-time signal. Future studies are of interest to optimally exploit all information in the four-dimensional time-reversal data.

B. Relevance vector machine (RVM) regression

Tipping²¹ has proposed the relevance vector machine classifier as a Bayesian generalization of the support vector machine (SVM).³⁶ Given a feature vector \mathbf{v} , our goal is to perform classification based on the probability that \mathbf{v} is associated with hypotheses H_0 (false target) and H_1 (true target), denoted respectively by $p(H_0|\mathbf{v})$ and $p(H_1|\mathbf{v})$. RVM-based regression is discussed first, which will set the stage for designing the classifier $p(H_1|\mathbf{v})$.

Assume access to “training” data $\{\mathbf{v}_n, l_n\}_{n=1,N}$, where \mathbf{v}_n represent feature vectors and l_n represent scalar labels. In the regression problem the objective is to learn a function that predicts the label of a given feature vector. Predictor function f can be written as

$$f(\mathbf{v}, \mathbf{w}) = \sum_{n=1}^N \mathbf{w}_n g(\mathbf{v}, \mathbf{v}_n) + \mathbf{w}_0, \quad (7)$$

where $g(\mathbf{v}, \mathbf{v}_n)$ is a kernel that quantifies the similarity between feature vectors \mathbf{v} and \mathbf{v}_n . Any linear or nonlinear measure may be selected for the kernel. The radial basis function (RBF) kernel^{21,23} is used in this paper where the similarity between two feature vectors \mathbf{v} and \mathbf{v}_n is defined as $g(\mathbf{v}, \mathbf{v}_n) = \exp[-\|\mathbf{v} - \mathbf{v}_n\|/2\sigma^2]$ with $\sigma=0.5$.

It is assumed that the error ϵ_n between the regression function $f(\mathbf{v}_n, \mathbf{w})$ and true label l_n is an independent, identically distributed (iid) zero-mean Gaussian with variance σ^2 , yielding $l_n = f(\mathbf{v}_n, \mathbf{w}) + \epsilon_n$, and for the N training examples the likelihood of the label set \mathbf{l} can be represented as

$$p(\mathbf{l}|\mathbf{w}, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\left(-\frac{\|\mathbf{l} - \Phi\mathbf{w}\|^2}{2\sigma^2}\right), \quad (8)$$

where $\mathbf{l} = \{l_1, \dots, l_N\}$, $\mathbf{w} = \{w_0, \dots, w_N\}$, and Φ is a kernel design matrix of size $N \times (N+1)$ whose i th row $\Phi_i = [1, g(\mathbf{v}_i, \mathbf{v}_1), \dots, g(\mathbf{v}_i, \mathbf{v}_N)]$. The weight w_n quantifies the importance of the training feature vector \mathbf{v}_n on the regression function $f(\mathbf{v}_n, \mathbf{w})$. The likelihood of label \mathbf{l} [see Eq. (8)] is dependent on the weights \mathbf{w} and variance σ^2 ; although not shown explicitly, the likelihood of \mathbf{l} is also dependent on the feature vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N\}$, via the kernel matrix Φ .

Our objective is to learn the weights \mathbf{w} of the regression model, based on the training data $\{\mathbf{v}_n, l_n\}_{n=1, \dots, N}$. A prior distribution is introduced over the weights \mathbf{w} , governed by a set of hyperparameters $\boldsymbol{\alpha}$, defining $p(\mathbf{w}|\boldsymbol{\alpha})$. The unknowns to be estimated are now the weights \mathbf{w} , the hyperparameters $\boldsymbol{\alpha}$, and the variance σ^2 . The introduction of hyperparameters $\boldsymbol{\alpha}$ appears counterintuitive, since the number of unknowns has essentially been doubled. However, careful design of $p(\mathbf{w}|\boldsymbol{\alpha})$ makes these additional unknowns desirable. Specifically, as discussed further below, the density function $p(\mathbf{w}|\boldsymbol{\alpha})$ and prior probability distribution on the hyperparameters $p(\boldsymbol{\alpha})$ are designed such that they constrain the properties of \mathbf{w} learned during the training phase, thereby serving as a regularization that avoids overfitting. Overfitting occurs when the weights \mathbf{w} match the training data $\{\mathbf{v}_n, l_n\}_{n=1, \dots, N}$ very well, but the regression model does not generalize well to feature vectors \mathbf{v} that are not in the training set.

The likelihood of the weights \mathbf{w} , hyperparameters $\boldsymbol{\alpha}$, and variance σ^2 given the training data is expressed as

$p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2 | \mathbf{I}) = p(\mathbf{w} | \boldsymbol{\alpha}, \sigma^2, \mathbf{I}) p(\boldsymbol{\alpha}, \sigma^2 | \mathbf{I})$. The first term on the right can be written as $p(\mathbf{w} | \boldsymbol{\alpha}, \sigma^2, \mathbf{I}) = p(\mathbf{I} | \mathbf{w}, \sigma^2) p(\mathbf{w} | \boldsymbol{\alpha}) / p(\mathbf{I} | \boldsymbol{\alpha}, \sigma^2)$. By choosing $p(\mathbf{w} | \boldsymbol{\alpha})$ as a product of Gaussians, $p(\mathbf{w} | \boldsymbol{\alpha}) = \prod_{n=0}^N N(w_n | 0, \alpha_n^{-1})$, the integral $p(\mathbf{I} | \boldsymbol{\alpha}, \sigma^2) = \int p(\mathbf{I} | \mathbf{w}, \sigma^2) p(\mathbf{w} | \boldsymbol{\alpha}) d\mathbf{w}$ can be evaluated analytically (it is a convolution of Gaussians). Now combining $p(\mathbf{I} | \boldsymbol{\alpha}, \sigma^2)$ with (8) yields

$$p(\mathbf{w} | \mathbf{I}, \boldsymbol{\alpha}, \sigma^2) = (2\pi)^{-(N+1)/2} |\boldsymbol{\Sigma}|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{w} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{w} - \boldsymbol{\mu})\right\}, \quad (9)$$

where the mean and covariance of the posterior distribution are given by

$$\boldsymbol{\Sigma} = (\sigma^{-2} \boldsymbol{\Phi}^T \boldsymbol{\Phi} + \mathbf{A})^{-1}, \quad \boldsymbol{\mu} = \sigma^{-2} \boldsymbol{\Sigma} \boldsymbol{\Phi}^T \mathbf{I} \quad (10)$$

with $\mathbf{A} = \text{diag}\{\alpha_0, \dots, \alpha_N\}$.

The evaluation of $p(\boldsymbol{\alpha}, \sigma^2 | \mathbf{I}) \propto p(\mathbf{I} | \boldsymbol{\alpha}, \sigma^2) p(\boldsymbol{\alpha}) p(\sigma^2)$ requires introduction of prior distributions $p(\boldsymbol{\alpha})$ and $p(\sigma^2)$. Since $\boldsymbol{\alpha}$ and σ^2 are associated with the variances of Gaussian distributions, they must be purely positive. Therefore gamma priors are introduced

$$p(\boldsymbol{\alpha}) = \prod_{n=0}^N \text{Gamma}(\alpha_n | a, b), \quad p(\beta) = \text{Gamma}(\beta | c, d) \quad (11)$$

with $\beta = \sigma^{-2}$ and $\text{Gamma}(\alpha_n | a, b) = \Gamma(a)^{-1} b^a \alpha_n^{a-1} e^{-b\alpha_n}$, with $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$. The parameters a, b, c , and d are set to zero, yielding a set of uniform hyperpriors (Jeffrey's prior²¹).

The expression $p(\boldsymbol{\alpha}, \sigma^2 | \mathbf{I})$ cannot be evaluated analytically, and therefore it is approximated as a delta function at its mode: $p(\boldsymbol{\alpha}, \sigma^2 | \mathbf{I}) \approx \delta(\boldsymbol{\alpha} - \boldsymbol{\alpha}_{MP}) \delta(\sigma^2 - \sigma_{MP}^2)$. The most-probable $\boldsymbol{\alpha}_{MP}$ and σ_{MP}^2 are computed by maximizing $p(\boldsymbol{\alpha}, \sigma^2 | \mathbf{I})$, which for the uniform hyperparameters reduces to maximizing

$$p(\mathbf{I} | \boldsymbol{\alpha}, \sigma^2) = \int p(\mathbf{I} | \mathbf{w}, \sigma^2) p(\mathbf{w} | \boldsymbol{\alpha}) d\mathbf{w} \\ = (2\pi)^{-N/2} |\sigma^2 \mathbf{I} + \boldsymbol{\Phi} \mathbf{A}^{-1} \boldsymbol{\Phi}^T|^{-1/2} \\ \times \exp\left\{-\frac{1}{2} \mathbf{I}^T (\sigma^2 \mathbf{I} + \boldsymbol{\Phi} \mathbf{A}^{-1} \boldsymbol{\Phi}^T)^{-1} \mathbf{I}\right\}. \quad (12)$$

The maximization yields

$$\alpha_n^{new} = \gamma_n / \mu_n^2, \quad \text{where} \quad \gamma_n \equiv 1 - \alpha_n \sum_{nn}, \quad (13)$$

and

$$(\sigma^2)^{new} = \frac{\|\mathbf{I} - \boldsymbol{\Phi} \boldsymbol{\mu}\|^2}{N - \sum_n \gamma_n}. \quad (14)$$

The learning process is therefore characterized by repeated application of (13) and (14), concurrent with updating $\boldsymbol{\Sigma}$ and $\boldsymbol{\mu}$ from (10).

The composite prior distribution on the weights, defined by $p(\mathbf{w} | \boldsymbol{\alpha}) = \prod_{n=0}^N N(w_n | 0, \alpha_n^{-1})$ and $p(\boldsymbol{\alpha}) = \prod_{n=0}^N \text{Gamma}(\alpha_n | a, b)$, yields an algorithm in which most of the α_n in (13) go to "infinity" (reaching the maximum precision of the computer), and the associated weight w_n has a likelihood highly peaked at zero. The associated feature vector \mathbf{v}_n is deemed "irrelevant" to the function $f(\mathbf{v}, \mathbf{w})$ that one wishes to learn, and the irrelevant feature vectors are

pruned via iterative update of Eqs. (13) and (14). The algorithm therefore works by seeking to maximize the likelihood $p(\boldsymbol{\alpha}, \sigma^2 | \mathbf{I})$, under the constraint imposed by the prior distributions $p(\mathbf{w} | \boldsymbol{\alpha})$ and $p(\boldsymbol{\alpha})$.

Once the most-probable parameters $\boldsymbol{\alpha}_{MP}$ and σ_{MP}^2 are estimated within the training phase, the joint posterior of model parameters is approximated as $p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2 | \mathbf{I}) = p(\mathbf{w} | \boldsymbol{\alpha}, \sigma^2, \mathbf{I}) p(\boldsymbol{\alpha}, \sigma^2 | \mathbf{I}) \approx p(\mathbf{w} | \boldsymbol{\alpha}_{MP}, \sigma_{MP}^2, \mathbf{I})$. For given weights \mathbf{w} the likelihood of label l_* being associated with a feature vector \mathbf{v}_* is expressed as, from (8), $p(l_* | \mathbf{v}_*, \mathbf{w}, \sigma_{MP}^2) = (2\pi \sigma_{MP}^2)^{-1/2} \exp(-\|l_* - \boldsymbol{\phi}(\mathbf{v}_*) \mathbf{w}\|^2 / 2\sigma_{MP}^2)$ where the vector $\boldsymbol{\phi}(\mathbf{v}_*)$ is defined as $\boldsymbol{\phi}(\mathbf{v}_*) = [1, g(\mathbf{v}_*, \mathbf{v}_1), \dots, g(\mathbf{v}_*, \mathbf{v}_N)]^T$. Given these expressions, the likelihood of target l_* for feature vector \mathbf{v}_* is expressed as

$$p(l_* | \mathbf{I}, \boldsymbol{\alpha}_{MP}, \sigma_{MP}^2) = \int p(l_* | \mathbf{v}_*, \mathbf{w}, \sigma_{MP}^2) p(\mathbf{w} | \boldsymbol{\alpha}_{MP}, \sigma_{MP}^2, \mathbf{I}) d\mathbf{w} = N(l_* | y_*, \sigma_*^2) \quad (15)$$

with

$$y_* = \boldsymbol{\mu}^T \boldsymbol{\phi}(\mathbf{v}_*), \quad \sigma_*^2 = \sigma_{MP}^2 + \boldsymbol{\phi}(\mathbf{v}_*)^T \boldsymbol{\Sigma} \boldsymbol{\phi}(\mathbf{v}_*). \quad (16)$$

The feature vector \mathbf{v}_* may therefore be mapped to $l_* = y_*$, and the uncertainty in this mapping is reflected by σ_*^2 .

C. Relevance vector machine (RVM) classifier

For the classification problem the labels l are binary, where $l=1$ corresponds to a target and $l=0$ to a nontarget (clutter). Introducing a logistic link²¹ as

$$\sigma[f(\mathbf{v}; \mathbf{w})] = 1 / \{1 + \exp[-f(\mathbf{v}; \mathbf{w})]\}, \quad (17)$$

the probability of labels \mathbf{I} (for the N training examples) may be expressed as

$$p(\mathbf{I} | \mathbf{w}) = \prod_{n=1}^N [\sigma\{f(\mathbf{v}_n; \mathbf{w})\}]^{I_n} [1 - \sigma\{f(\mathbf{v}_n; \mathbf{w})\}]^{1-I_n}. \quad (18)$$

The model-parameter posterior distribution can be written as $p(\mathbf{w}, \boldsymbol{\alpha} | \mathbf{I}) = p(\mathbf{w} | \boldsymbol{\alpha}, \mathbf{I}) p(\boldsymbol{\alpha} | \mathbf{I})$, where the regression variance σ^2 is no longer present [note that the use of the $\sigma(\cdot)$ to represent the logistic link function is widely employed within the statistics community, and this must be distinguished from the variance σ^2]. The approximation $p(\boldsymbol{\alpha} | \mathbf{I}) \approx \delta(\boldsymbol{\alpha} - \boldsymbol{\alpha}_{MP})$ is again employed, using the same sparseness prior distributions $p(\mathbf{w} | \boldsymbol{\alpha})$ and $p(\boldsymbol{\alpha})$ as employed in the regression problem. The computation of $p(\mathbf{w}, \boldsymbol{\alpha} | \mathbf{I}) = p(\mathbf{w} | \boldsymbol{\alpha}, \mathbf{I}) p(\boldsymbol{\alpha} | \mathbf{I})$ depends on the evaluation of $p(\mathbf{w} | \boldsymbol{\alpha}, \mathbf{I})$, which for the classification problem is not Gaussian (as it was for the regression problem). In the interest of obtaining a solution analogous to that found for the regression formulation, $p(\mathbf{w} | \boldsymbol{\alpha}, \mathbf{I})$ is approximated by a Gaussian distribution. The associated mean is computed as the maximum of $p(\mathbf{w} | \boldsymbol{\alpha}, \mathbf{I}) \propto p(\mathbf{I} | \mathbf{w}) p(\mathbf{w} | \boldsymbol{\alpha})$, with the covariance matrix represented by the Hessian of $p(\mathbf{w} | \boldsymbol{\alpha}, \mathbf{I}) \propto p(\mathbf{I} | \mathbf{w}) p(\mathbf{w} | \boldsymbol{\alpha})$; this yields the so-called Laplace approximation.²¹ The most-probable weights \mathbf{w}_{MP} can be computed analytically, as can the Hessian. Once the Laplace approximation is so employed, the training of the classifier is performed by an iterative algorithm, exactly as developed above for RVM regres-

sion. During this training procedure the most-probable weights \mathbf{w}_{MP} are characterized by many \mathbf{w}_n being set to zero, with only the most-relevant training examples (those with nonzero weights) used in the final classifier.

Now consider a feature vector \mathbf{v} that needs to be classified. The probability that \mathbf{v} is associated with label $l=1$ (corresponding to the probability that \mathbf{v} corresponds to a target of interest) is evaluated by computing (17) using the most-probable weights \mathbf{w}_{MP} , determined when training. This yields the desired $p(H_1|\mathbf{v})$.

V. TARGET CLASSIFICATION PERFORMANCE

A. Problem details

To consider changes in the channel parameters, the channel depth is varied between 6 and 9 m, with all other channel parameters held constant. In each test case, it is assumed that a submerged target has already been detected and our objective is to determine if it is one of the targets seen when training (or if it is a clutter item). All examples shown in this paper correspond to simulated targets where the free-field bistatic responses are computed using an FEM code³¹ and the in-channel response is computed using the free-field response and normal-mode channel model (KRAKEN²⁹). The performance of time-reversal imaging has been considered by us previously using measured data;¹ simulated data are considered here such that a range of uncertainty in the channel parameters may be examined, as well as a range of target/clutter types (the measured data in Ref. 1 is too limited to perform the classification studies of interest here). In each test case, the channel parameters are estimated via GS, before TRI is performed. Five distinct scatterers are placed within the acoustic channel. Two of the targets are deemed “targets of interest” (defined by hypothesis one, H_1), and the other three are deemed “false targets” (defined by null hypothesis, H_0). The incident fields on the target propagate through the channel from an unidirectional source, located at the center of the sensor array. The scattered fields are computed for observations in a vertical plane, this plane situated

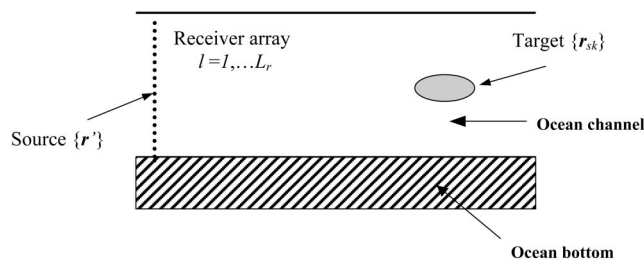


FIG. 1. Schematic of array-based sensing of a target in an acoustic waveguide. A single isotropic source (\mathbf{r}') is located at the center of the receiver array, and \mathbf{r}_{rl} , $rl=1,\dots,L_r$, locate the array elements. The k th target scattering center is located at \mathbf{r}_{sk} . The position of the receiver array and depth of the target relative to the bottom are kept fixed, and the target is rotated relative to the array to generate multiple target-sensor orientations.

within the channel, at a fixed distance from the target center. The fields within the plane are then coupled to a modal channel model,²⁹ with which they are propagated away from the target through the channel to the receiver array. The target-receiver distance is fixed at 100 m. In addition, the center of the target is 0.12 m above the channel bottom. A 28-element linear receiver array is employed perpendicular to the bottom (see Fig. 1). The ocean bottom is treated as a smooth horizontal surface to which the array is vertical. The array is 2 m long and its center is situated 2 m below the sea surface. A two-layer geoacoustic model is considered, with sound speed in water of 1520 m/s, a bottom density of 1.5 g/cm³, a channel-bottom attenuation factor of 0.28 dB/KmHz, and sound speed in the bottom of 1800 m/s for simulating the in-channel response at the receiver array.

The two scatterers in H_1 are minelike targets (a steel and fiberglass cylinder). The three false targets are a 55-gallon drum, an oxygen tank, and a rock. The data covers the frequency band 50 Hz to 10 KHz, and the scattered waveforms received by the vertical receiver array are considered in the time domain (after Fourier synthesis), as propagated through the channel to the vertical receiver array. Each of the five targets is rotated about its axis over 180°, with 1° angular

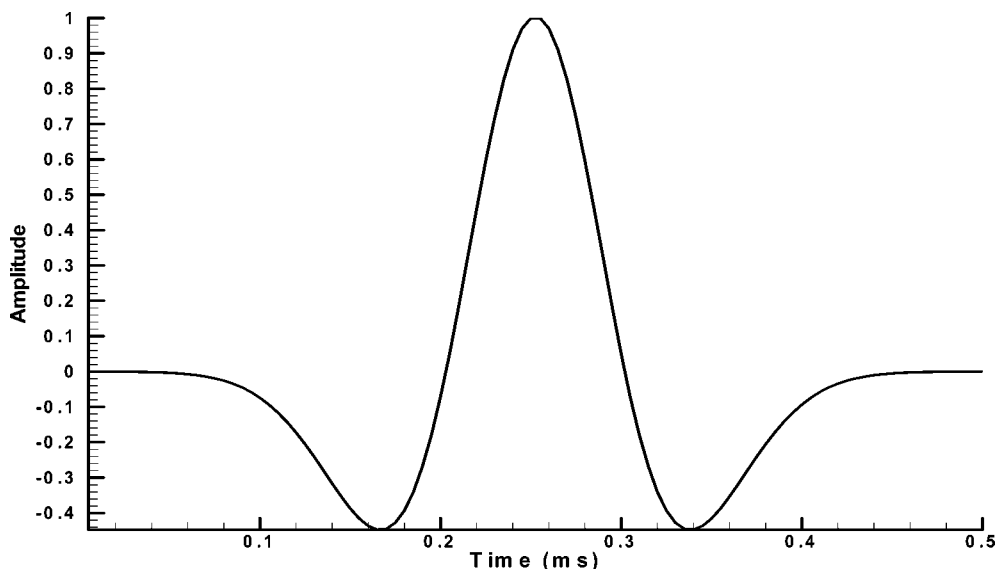


FIG. 2. Incident pulse shape for the synthesized time-domain scattering data.

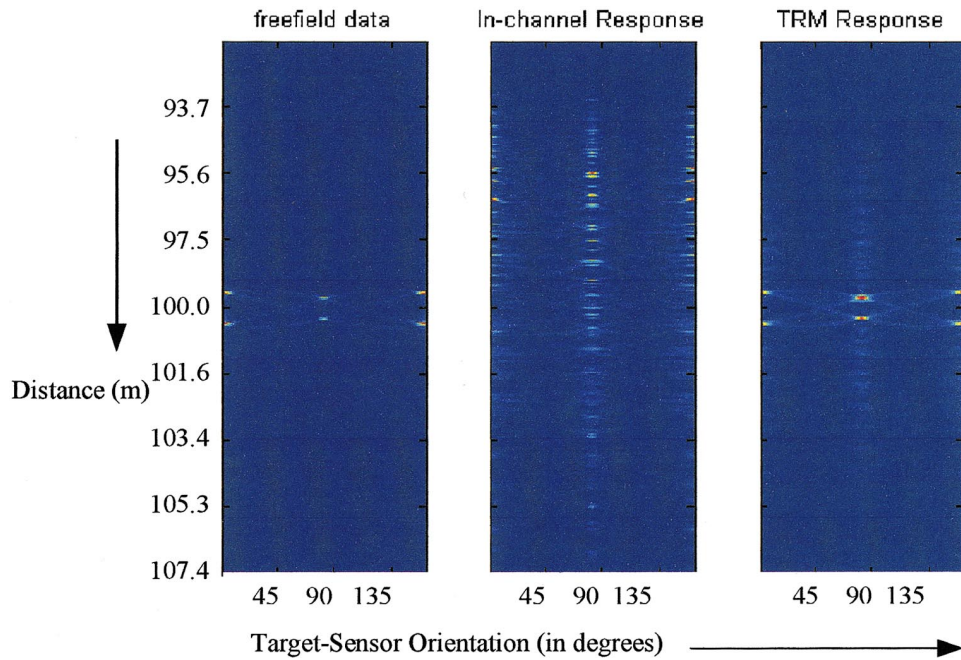


FIG. 3. Scattering data from a 55-gallon drum in the channel. The horizontal axis denotes the orientation of the target relative to the vertical sensor array (broadside or specular reflection at 90°) and the vertical axis represents target-sensor distance in meters (which may be related to time, via the sound speed). The drum is on the channel bottom, with center 0.12 m above the channel floor at a range of 100 m (6-m channel depth): (a) synthesized free-field time-domain data, (b) scattering data as received by the center element of the receiver array, and (c) time-reversed signal through the target center, using channel parameters estimated via Gibbs sampling. Each subimage is normalized such that dynamic range is between 0 and 1.

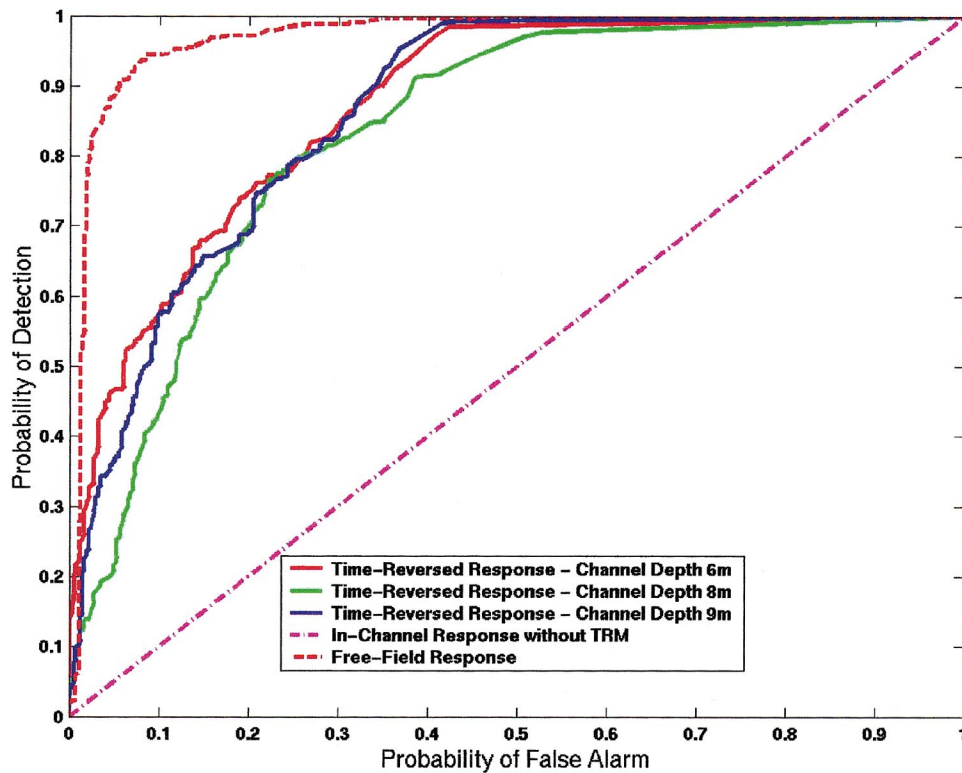


FIG. 4. Performance of the RVM classifier, trained using free-field signatures from two targets of interest (H_1). The dashed line shows the classification performance when tested on free-field responses from all targets and the dotted line shows the performance of the same classifier when tested directly on in-channel responses of all targets (at all orientations and at depths of 6, 8, and 9 m). The solid lines in different colors show the performance when time-reversed in-channel target signatures were subjected to the RVM classifier. Different colors represent different channel depths. Gibbs sampling was used to estimate the channel parameter prior to time-reversal. The data included three “false targets” not observed while training. Target-sensor distance was fixed at 100 m.

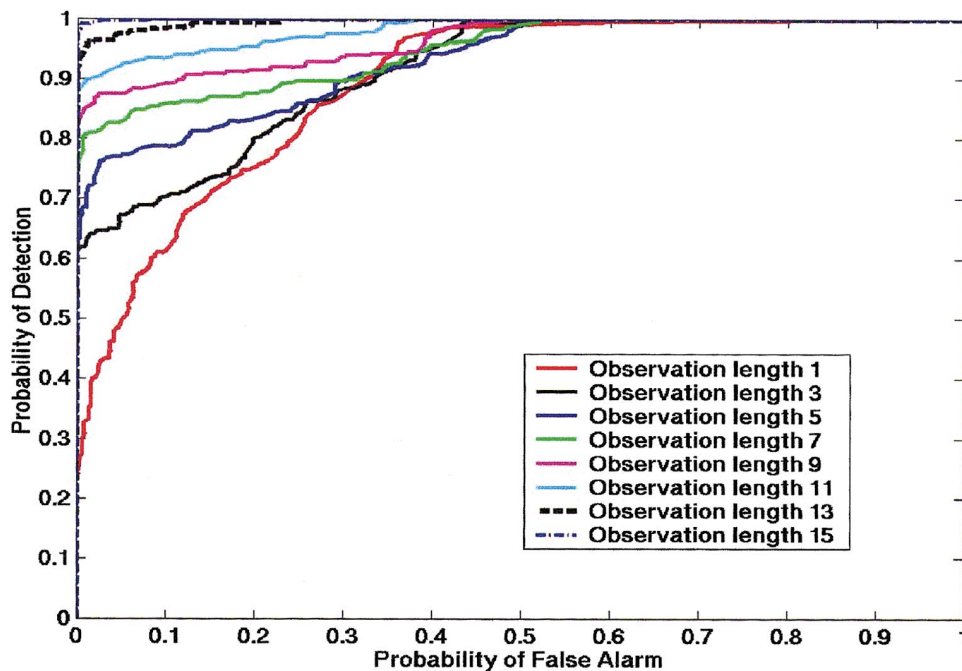


FIG. 5. The classification performance of the RVM classifier, trained on free-field response of two minelike targets and tested on multi-aspect scattering data sampled in 5° increments. The figure shows classification performance for different sequence lengths. For all sequences the angular increment between successive observations is 5° . The relative change in target-sensor orientation is assumed known, but the absolute target-sensor orientation is unknown to the classifier.

sampling, thereby constituting multiple synthesized measurements. The shape of the pulse driving the point-source transmitter is shown in Fig. 2.

B. Geoacoustic inversion via Gibbs sampling

The parameters of the geoacoustic model (e.g., depth, sound speed in the water and bottom, channel depth, etc.) are assumed unknown at the receiver array although the target location is assumed known (100 m along downrange and 0.12 m above the floor). Note that our objective is to perform target classification, and it is assumed that detection has already been achieved. The target depth and position could also be accounted within the GS procedure, as has been done in a previous genetic algorithm study.¹ The channel parameters considered for GS estimation are channel depth, sound speed in water, sound speed in the bottom, and the density and attenuation of the bottom. Since GS tries to address channel estimation in a Bayesian construct, one needs to define the model space and the *a priori* probability of the model parameters within that space. The model space is defined by channel depth varying between 5 and 10 m, sound speed in water between 1470 and 1550 m/s, sound speed of the bottom between 1750 and 1850 m/s, bottom attenuation between 0.25 and 0.35 dB/Km Hz, and density of the bottom between 1 and 2 g/cm³. In our problem a uniform *a priori* probability is assumed within the model space. GS is employed for a single frequency (the frequency that corresponds to the strongest signal at the receiver array, 2.25 KHz) although the algorithm is capable of handling multiple frequencies, as presented in (6). For noise-free data the marginal posterior distribution of individual channel parameters is observed to be approximately a delta function around the

true values; details for noisy data are given below. The comparison between the true and estimated channel parameters is shown in Table I for scattering from an oil-drum situated in a 6-m-deep channel, for noise-free data. The GS-based geoacoustic inversion is performed for each target separately and the MAP estimates are used from the *a posteriori* parameter density for subsequent time-reversal imaging.

To demonstrate the form of the data, Fig. 3(a) represents the free-field scattered response from the 55-gallon drum. The vertical axis represents time and the horizontal axis denotes the target orientation. Scattering from the same target is shown in Fig. 3(b), situated in a 6-m-deep channel and 100 m from the receiver array, as received by the center receiver element. It is important to note that the extended temporal support of the time-domain signal is reflective of the waveguide dispersion.

C. Time-reversal imaging and feature extraction

The time-reversed response of the 55-gallon drum is shown in Fig. 3(c). The vertical axis in Fig. 3(c) represents target-sensor range and the horizontal axis represents orientation of the drum relative to the receiver array. It is apparent from Fig. 3(c) that the time-reversal processing with an estimated channel model largely removes the channel-induced dispersion and recovers a signal similar to the free-field response [Fig. 3(a)]. The comparison between Figs. 3(a) and 3(c) motivates RVM design based on free-field data, with time-reversal employed to approximately remove channel effects. The time-reversed response for each target/clutter at every orientation is subjected to feature extraction and the feature vector is used for subsequent classification by the RVM.

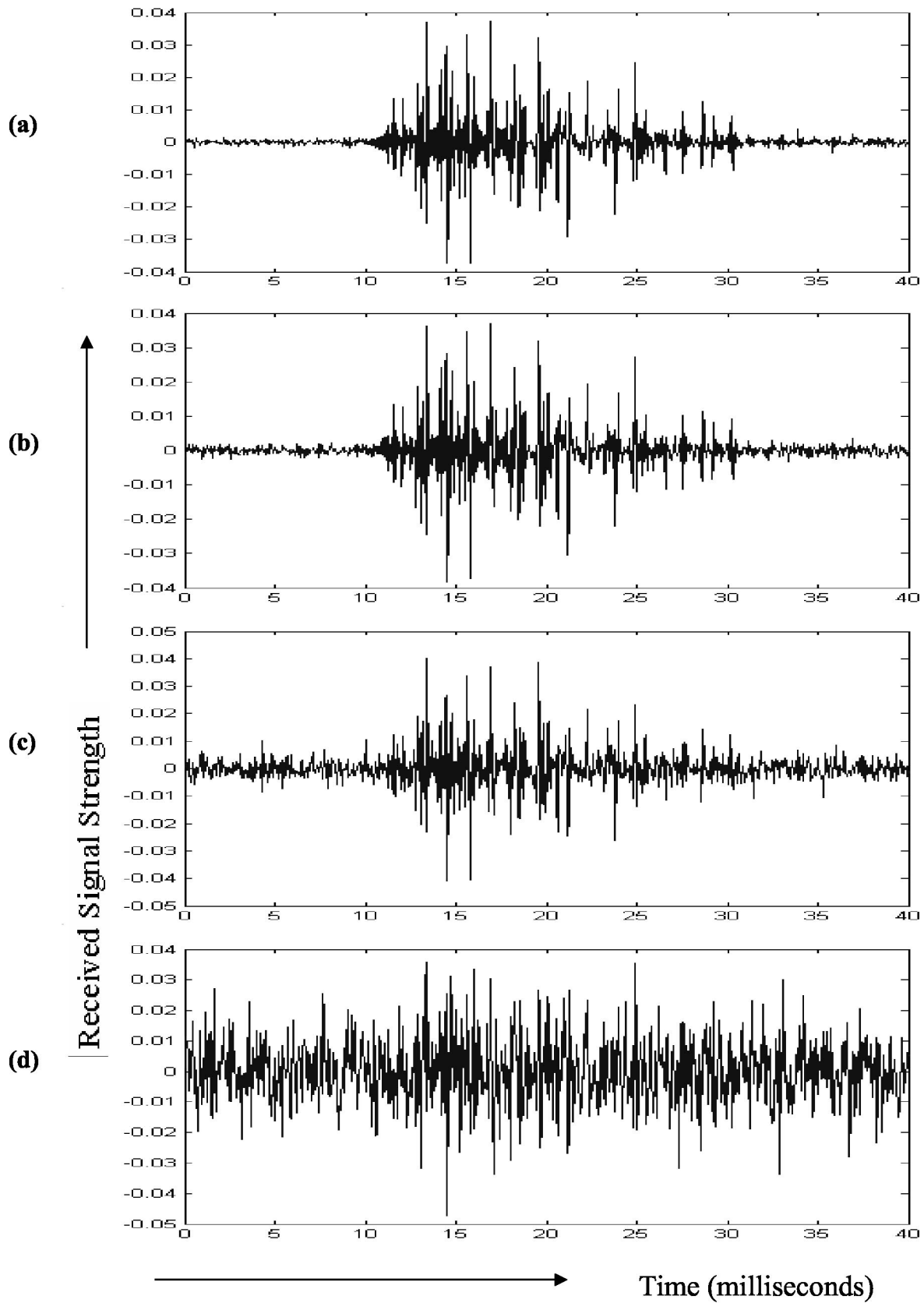


FIG. 6. Example scattered waveforms at various SNR levels. The additive noise is formed by convolving the incident-pulse shape (Fig. 2) with white Gaussian noise. These example results correspond to a 55-gallon drum situated in a 9-m-deep channel, with center 0.12 m above the floor, at a range of 100 m. These examples correspond to data received at the center element of the vertical array. The target is oriented at 90° with respect to the vertical array, corresponding to broadside excitation (strongest target signature): (a) noise-free data, (b) 10 dB SNR, (c) 0 dB SNR, and (d) -10 dB SNR. The horizontal axis for all the subplots is time in milliseconds.

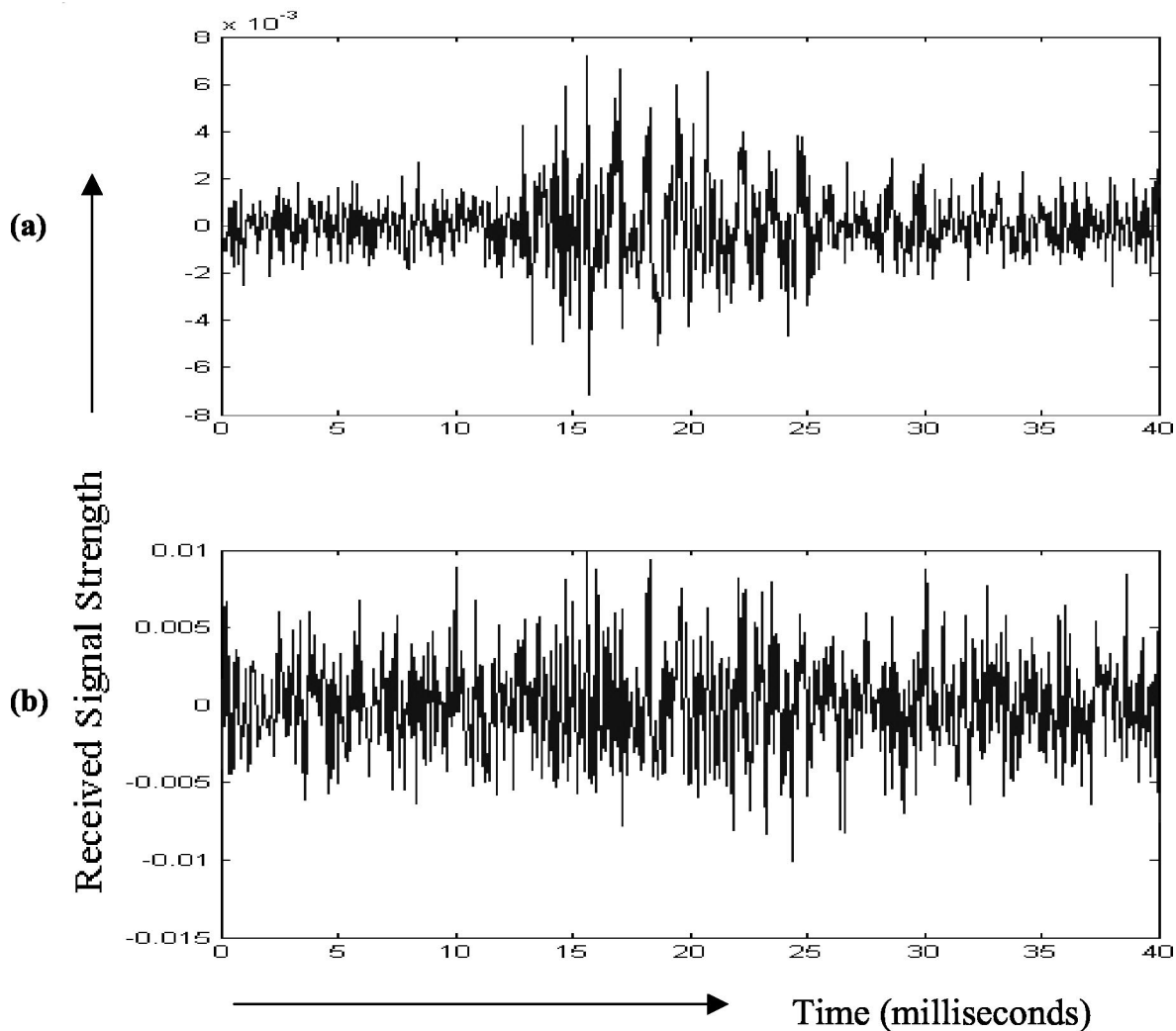


FIG. 7. As in Fig. 6, for the target-sensor orientation corresponding to weakest target-signature strength. The SNR values correspond to average target signature strength across all possible target-sensor orientations: (a) 10 dB SNR and (b) 0 dB SNR. The horizontal axis for both the subplots is time in milliseconds.

D. Classification results

As indicated above, a time-reversal signature is obtained for each orientation of the vertical array relative to the target, using the speed of sound to convert from the range-dependent time-reversal data. Features are extracted from this time-domain waveform. It is assumed that in each case the presence of a target or non-target has been detected, and our objective is to classify it. It needs to be emphasized that the RVM classifier²¹ is trained solely on free-field data from the two targets of interest (H_1), with no *a priori* knowledge assumed for the false targets (H_0) and with channel effects removed approximately via time reversal.

Figure 4 demonstrates the performance of the RVM classifier in terms of the receiver operating characteristic (ROC), quantified in terms of the probability of detection (probability that a target from H_1 is properly so classified) versus the probability of false alarm (probability that one of the three false targets is classified as being in H_1). The solid lines in three different colors depict the performances when the clas-

sifier is tested on the time-reversed responses for all targets and false targets, at channel depths of 6, 8, and 9 m. These results are based on classification at a single target-sensor orientation (with performance averaged across all orientations). The dash-dotted magenta line in Fig. 4 shows the RVM performance of the same classifier when tested directly on in-channel target data (no time-reversal), as received at the center element of the receiver array, using the same features as discussed above. The dashed red line in Fig. 4 shows the performance of the same RVM classifier when tested on the free-field response of all targets and non-targets (no channel). This represents the optimal performance that could be achieved if time-reversal removes the channel effects completely.

It has been demonstrated by several researchers^{37,38} that the performance of target classification may be enhanced by combining the information gathered from multiple target orientations (aspects). Multi-aspect classification is performed here as follows. Assume that vertical array data are collected

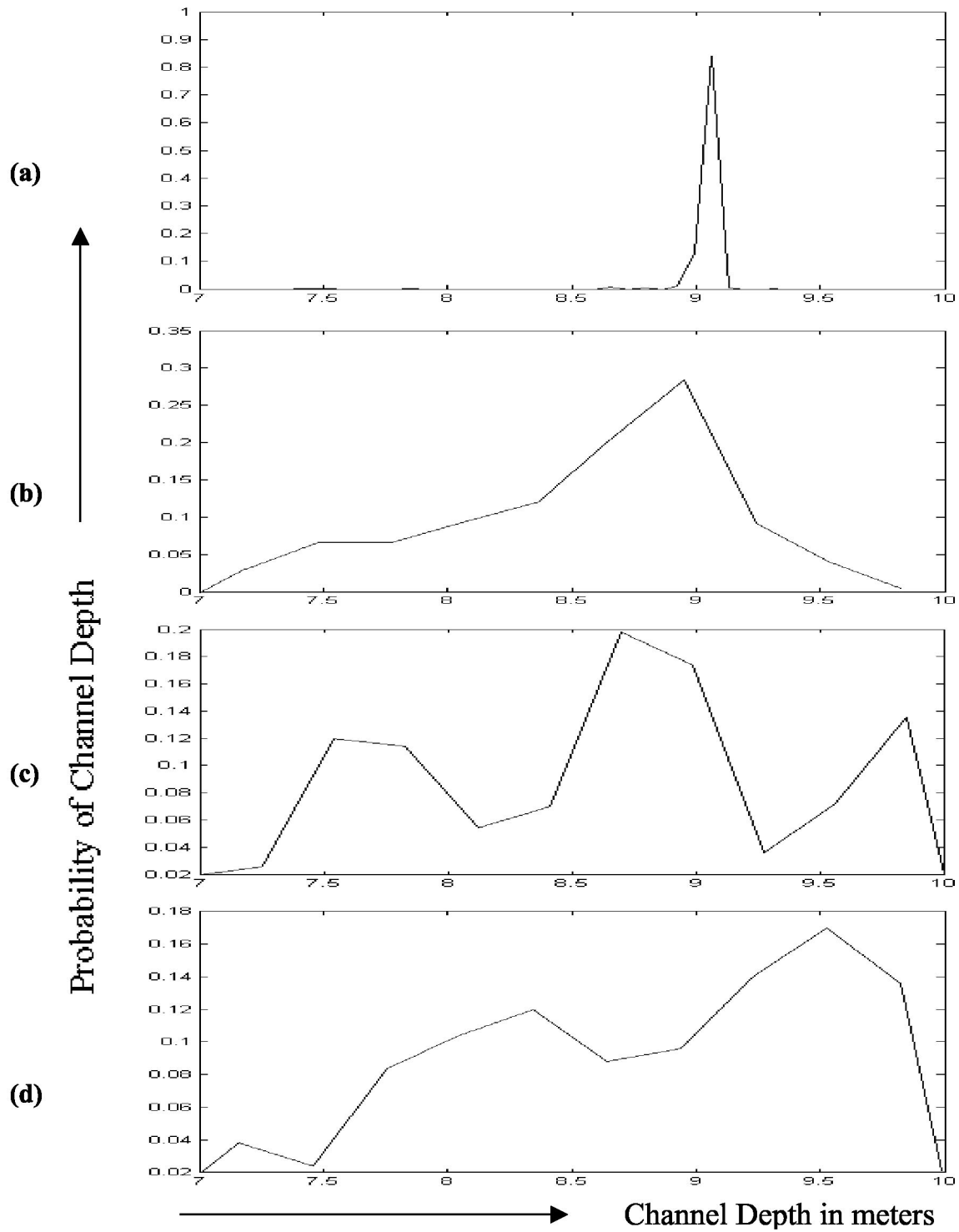


FIG. 8. The marginal posterior distribution of channel depth at different SNR levels (colored noise), over the 7–10-m search range. The true channel depth is 9 m: (a) 0 dB, (b) –5 dB, (c) –10 dB, and (d) –20 dB.

at multiple orientations (with respect to the target) along a circular path, constituting a synthetic aperture. Let \mathbf{x}_q represent the vertical-array data, for the q th position along the synthetic aperture. Given a set of backscattered responses from multiple target-sensor orientations, the response \mathbf{x}_q that corresponds to maximum signal strength is selected and the

channel parameters are estimated from \mathbf{x}_q via GS. Time-reversal imaging of \mathbf{x}_q ($\forall q=1, \dots, Q$) based on estimated channel parameters yields the associated feature vector, \mathbf{v}_q (Q feature vectors, one for each of the Q vertical-array positions, at multiple orientations about the target). Note that the

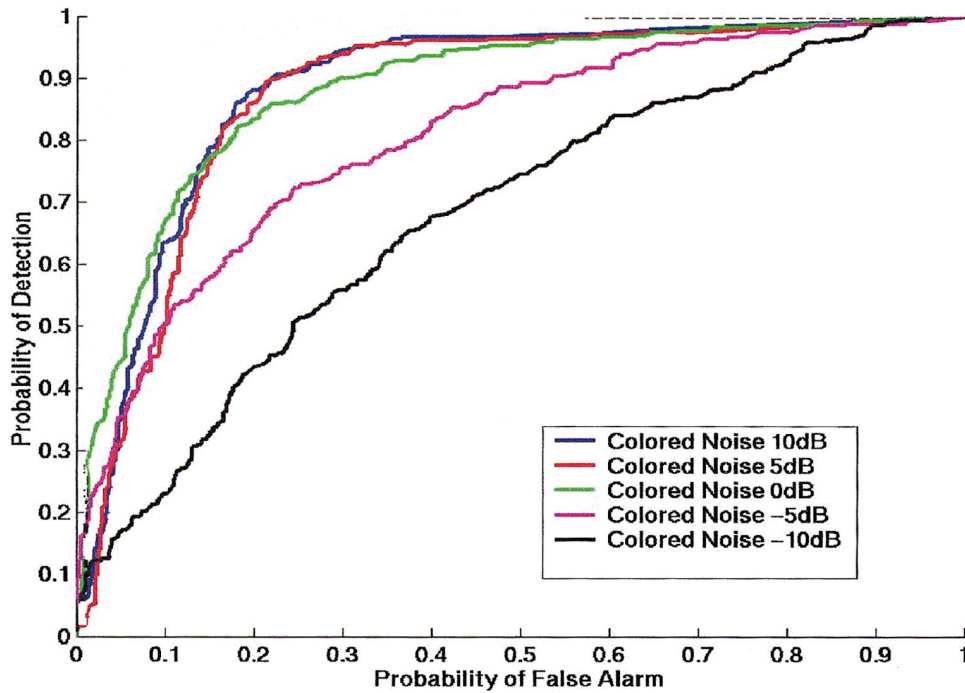


FIG. 9. Target classification performance for the case of additive colored noise. The colored noise was added to the backscattered response received by the sensor array before channel parameters were estimated.

channel parameters are estimated from the most significant backscattered response, not individually from each target-sensor orientation. The classifier in Eq. (17) quantifies the probability $p(H_1|\mathbf{v}_q)$ that the q th measurement is associated with class H_1 . For Q sensor positions along the synthetic aperture, the cumulative statistic is obtained as $P_Q = \prod_{q=1}^Q p(H_1|\mathbf{v}_q)$. Note that the multi-aspect waveforms are treated independently here for simplicity, although more-sophisticated techniques are available.³⁷ Synthetic apertures of multiple lengths are considered in Fig. 5, and, in each case, consecutive positions along the aperture correspond to changes of 5° in the target-sensor angle (angular sampling of 5°). As expected, the classification performance increases monotonically when observations from multiple aspects are fused.

E. Noisy data

All examples considered above assumed no noise at the receiver array. In the following examples noise is added at the receiver array before estimation of the channel parameters. To approximate scattering from rough channel interfaces (top and bottom), the additive colored noise was synthesized by convolving white Gaussian noise with the incident-pulse shape. The noisy time-domain waveforms at the receiver were converted to the frequency domain and subjected to channel-parameter estimation. Figure 6(a) shows the noise-free in-channel response for a 55-gallon drum situated in a 9-m-deep water channel, with center 0.12 m above the channel bottom, as received by a single receiver in the array center. The target response corresponds to a target-sensor orientation of 90° . Figures 6(b)–6(d) show the in-channel drum response at noise levels of 10, 0, and -10 dB, respectively, defined in terms of the signal energy divided by the noise energy. For a given value of SNR, the noise variance is the same for each target-sensor orientation considered. Since the signal strength is a strong function of

orientation, the SNR is in fact target-sensor-orientation dependent. The SNR values mentioned above represent average values across all orientations. Figures 7(a) and 7(b) show the noisy in-channel response at 45° orientation for the same drum, for noise variances associated with average SNR values of 10 and 0 dB, respectively. A comparison of Figs. 6 and 7 underscores that the strength of the target-signature relative to the noise is a strong function of orientation.

As discussed above, at high average SNR values (>0 dB), the marginal posterior distributions for the channel parameters are strongly peaked. Figure 8(a) shows the marginal posterior distribution of channel depth for a 55-gallon oil drum at a SNR of 0 dB. It is apparent that the posterior distribution is strongly peaked around the true channel depth of 9 m. Similar characteristics of the posterior distribution are observed for other channel parameters, such as the sound speed in the water and the bottom, not shown here for brevity.

The effect of noise on channel parameter estimation via GS is portrayed in Fig. 8. The noisy in-channel response corresponds to the same target (55-gallon oil drum situated in a 9-m-deep channel), with all results for the orientation corresponding to maximum signal strength. As expected, the posterior distribution (along with the MAP estimate of channel depth) deviates from being strongly peaked as the noise level increases. Figures 8(b)–8(d) represent the marginal posterior distribution of channel depth at three levels of SNR: -5 , -10 , and -20 dB, respectively.

RVM classification performance is now considered for the noisy data (as in Fig. 4, wherein noise-free data were considered). Figure 9 shows the performance of the same RVM classifier as in Fig. 4 (trained on noise-free, free-field data), applied to noisy time-reversed target responses at different noise levels. These results are based on a single observation (*i.e.*, observation sequence length, $K=1$). The results are averaged across all possible target-sensor orientations,

and it demonstrates the expected vitiation in performance as SNR is diminished.

VI. CONCLUSIONS

Time-reversal imaging has been employed to remove channel-induced dispersion effects, associated with backscattered signatures from targets situated in a shallow channel. The channel parameters required of the time-reversal imager have been estimated within a Bayesian framework, via Gibbs sampling. The key new development is examination of target classification, based on the time-reversed image. The objective was to use time reversal to remove channel-induced effects, such that the same classifier [trained solely on the free-field response of the target(s)] may be employed for different channel parameters (here multiple channel depths have been considered). In this study an RVM²¹ classifier is employed, and it was found that the classification performance (after time reversal) was relatively invariant to channel parameters (across a large range of channel depths).

The features employed in the RVM classifier were based on the spectral properties of a one-dimensional waveform, computed by taking a one-dimensional “cut” in the full time-reversal image. In general, the time-reversal imager yields a four-dimensional image (space-time),³⁹ and in future research it is of interest to develop features that fully exploit this information base. In addition, all of the synthesized data corresponded to the case of a perfect (flat-surface) acoustic waveguide. Colored noise was added to the data, to synthesize scattering from a rough channel surface and bottom. However, future studies are needed to explicitly assess the effects of surface roughness on time-reversal-image quality, and on the subsequent classification performance.

- ¹L. Carin, H. Liu, T. Yoder, L. Couchman, B. Houston, and J. Bucaro, “Wideband time-reversal imaging of an elastic target in an acoustic waveguide,” *J. Acoust. Soc. Am.* **115**, 259–268 (2004).
- ²P. Gerstoft, “Ocean acoustic inversion with estimation of a posteriori probability distributions,” *J. Acoust. Soc. Am.* **104**, 808–819 (1998).
- ³R. P. Gorman and T. J. Sejnowski, “Learned classification of sonar targets using a massively parallel network,” *IEEE Trans. Acoust., Speech, Signal Process.* **36**, 1135–1140 (1988).
- ⁴G. Goo and W. L. Au, “Detection and identification of buried objects in shallow water,” in *Proc. SPIE Int. Symp. Aerospace/Defense Sensing Control*, Orlando, FL, April 1996, pp. 201–214.
- ⁵M. K. Broadhead, “Broadband source signature extraction from underwater acoustics data with sparse environment information,” *J. Acoust. Soc. Am.* **97**, 1322–1325 (1995).
- ⁶M. K. Broadhead, L. A. Pflug, and R. L. Field, “Use of high order statistics in source signature estimation,” *J. Acoust. Soc. Am.* **107**, 2576–2585 (2000).
- ⁷P. Runkle, P. Bharadwaj, and L. Carin, “Hidden Markov Models for multi-aspect target classification,” *IEEE Trans. Signal Process.* **47**, 2035–2040 (1999).
- ⁸W. A. Kuperman, W. S. Hodgkiss, H. C. Song, T. Akal, C. Ferla, and D. R. Jackson, “Phase conjugation in the ocean: Experimental demonstration of an acoustic time reversal mirror,” *J. Acoust. Soc. Am.* **103**, 25–40 (1998).
- ⁹A. Bageroer, W. A. Kuperman, and P. N. Mikhalevsky, “An overview of matched field methods in ocean acoustics,” *IEEE J. Ocean. Eng.* **18**(4), 401–424 (1993).
- ¹⁰P. Blomberg, G. Papanicolaou, and H. K. Zhao, “Super-resolution in time-reversal acoustics,” *J. Acoust. Soc. Am.* **111**, 230–248 (2002).
- ¹¹J. S. Kim, H. C. Song, and W. A. Kuperman, “Adaptive time-reversal mirror,” *J. Acoust. Soc. Am.* **109**, 1817–1825 (2001).
- ¹²D. R. Jackson and D. R. Dowling, “Phase conjugation in underwater acoustics,” *J. Acoust. Soc. Am.* **89**, 171–181 (1991).

- ¹³M. Fink, “Time-reversal acoustics,” *Phys. Today* **50**, 34–40 (1997).
- ¹⁴P. Stoica and A. Nehorai, “MUSIC, maximum likelihood, and Cramer-Rao Bound,” *IEEE Trans. Acoust., Speech, Signal Process.* **37**(5), 720–741 (1989).
- ¹⁵S. K. Lehman and A. J. Devaney, “Transmission mode acoustic time-reversal imaging for nondestructive evaluation,” *J. Acoust. Soc. Am.* **112**(5), 2390–2400 (2002).
- ¹⁶H. Liu, N. Dasgupta, and L. Carin, “Time-reversal imaging for wideband underwater target classification,” in *Proceedings of the International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Hong Kong, 6–10 April, 2003.
- ¹⁷S. E. Dosso, “Quantifying uncertainty in geoacoustic inversion. A fast Gibbs sampler approach,” *J. Acoust. Soc. Am.* **111**, 129–142 (2002).
- ¹⁸S. E. Dosso, “Quantifying uncertainty in geoacoustic inversion. Application to broadband, shallow-water data,” *J. Acoust. Soc. Am.* **111**, 143–159 (2002).
- ¹⁹M. K. Sen and P. L. Stoffa, “Bayesian inference, Gibbs’ sampler and uncertainty estimation in geophysical inversion,” *Geophys. Prospect.* **44**, 313–350 (1996).
- ²⁰N. Dasgupta and L. Carin, “Time-reversal imaging of distant targets in a shallow water channel,” *Proc. Oceans 2003*, San Diego, 22–26 September, 2003.
- ²¹M. E. Tipping, “Sparse Bayesian learning and the relevance vector machine,” *J. Mach. Learn. Res.* **1**, 211–244 (2001).
- ²²N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machine* (Cambridge U.P., Cambridge, 2000).
- ²³S. A. Solla, T. K. Leen, and K.-R. Müller (Ed.), “The Relevance Vector Machine,” in *Advances in Neural Information Processing Systems 12* (MIT, Cambridge, MA, 2001), pp. 652–658.
- ²⁴P. Roux and M. Fink, “Time reversal in a waveguide: Study of the temporal and spatial focusing,” *J. Acoust. Soc. Am.* **107**, 2418–2429 (2000).
- ²⁵L. Borcea, G. Papanicolaou, C. Tsogka, and J. Berryman, “Imaging and time reversal in random media,” *Inverse Probl.* **18**, 1247–1279 (2002).
- ²⁶C. Borden and M. Fink, “Mathematical foundations of the time reversal mirror,” *Asymptotic Anal.* **29**(2), 157–182 (2002).
- ²⁷N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, “Equation of state calculation by fast computing machines,” *J. Chem. Phys.* **21**, 1087–1092 (1953).
- ²⁸A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis* (Chapman & Hall, London, 1995).
- ²⁹M. B. Porter, “The KRAKEN Normal Mode Program,” http://oalib.saic.com/Modes/AcousticsToolbox/manual_html/kraken.html. See also M. B. Porter, “A numerical method for computing ocean acoustic modes,” Ph.D. dissertation, Northwestern Univ., 1984.
- ³⁰F. Jensen, W. Kuperman, M. Porter, and H. Schmidt, *Computational Ocean Acoustics*, (Springer-Verlag, New York, 2000).
- ³¹B. Szabo and I. Babuska, *Finite Element Analysis* (Wiley, New York, 1991).
- ³²J. J. Grannell, J. J. Shirron, and L. S. Couchman, “A hierarchic p-version boundary-element method for axisymmetric acoustic scattering and radiation,” *J. Acoust. Soc. Am.* **95**, 2320–2329 (1994).
- ³³B. H. Houston, M. H. Marcus, J. A. Bucaro, E. G. Williams, and D. M. Photiadis, “The structural acoustics and active control of interior noise in a ribbed cylindrical shell,” *J. Acoust. Soc. Am.* **99**, 3497–3512 (1996).
- ³⁴F. K. Gruber, E. A. Marengo, and A. J. Devaney, “Time-reversal imaging with multiple signal classification considering multiple scattering between targets,” *J. Acoust. Soc. Am.* **115**, 3042–3047 (2004).
- ³⁵L. Borcea, G. Papanicolaou, and C. Tsogka, “Theory and applications of time reversal and interferometric imaging,” *Inverse Probl.* **19**, 5139–5164 (2003).
- ³⁶J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis* (Cambridge U.P., Cambridge, 2004).
- ³⁷N. Dasgupta, P. Runkle, L. Couchman, and L. Carin, “Dual Hidden Markov Model for characterizing wavelet coefficients from multi-aspect scattering data,” *Signal Process.* **81**, 1303–1316 (2001).
- ³⁸N. Dasgupta, P. Runkle, L. Carin, L. Couchman, T. Yoder, J. Bucaro, and G. Dobeck, “Class-based Target Identification with Multi-Aspect Scattering Data,” *IEEE J. Ocean. Eng.* **28**(2), 271–282 (2003).
- ³⁹S. Kim, G. F. Edelmann, W. A. Kuperman, W. S. Hodgkiss, and H. C. Song, “Spatial resolution of time-reversal arrays in shallow water,” *J. Acoust. Soc. Am.* **110**, 820–829 (2001).

Protocols for calibrating multibeam sonar

Kenneth G. Foote, Dezhang Chu, and Terence R. Hammar
Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543

Kenneth C. Baldwin and Larry A. Mayer
University of New Hampshire, Durham, New Hampshire 03824

Lawrence C. Hufnagle, Jr.
Northwest Fisheries Science Center, NOAA-NMFS, Seattle, Washington 98112-2097

J. Michael Jech
Northeast Fisheries Science Center, NOAA-NMFS, Woods Hole, Massachusetts 02543

(Received 16 July 2004; revised 13 January 2005; accepted 20 January 2005)

Development of protocols for calibrating multibeam sonar by means of the standard-target method is documented. Particular systems used in the development work included three that provide the water-column signals, namely the SIMRAD SM2000/90- and 200-kHz sonars and RESON SeaBat 8101 sonar, with operating frequency of 240 kHz. Two facilities were instrumented specifically for the work: a sea well at the Woods Hole Oceanographic Institution and a large, indoor freshwater tank at the University of New Hampshire. Methods for measuring the transfer characteristics of each sonar, with transducers attached, are described and illustrated with measurement results. The principal results, however, are the protocols themselves. These are elaborated for positioning the target, choosing the receiver gain function, quantifying the system stability, mapping the directionality in the plane of the receiving array and in the plane normal to the central axis, measuring the directionality of individual beams, and measuring the nearfield response. General preparations for calibrating multibeam sonars and a method for measuring the receiver response electronically are outlined. Advantages of multibeam sonar calibration and outstanding problems, such as that of validation of the performance of multibeam sonars as configured for use, are mentioned. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1869073]

PACS numbers: 43.30.Xm, 43.58.Vb, 43.30.Yj, 43.30.Pc [WMC]

Pages: 2013–2027

I. INTRODUCTION

Multibeam sonar, with its simultaneous formation of multiple beams oriented differently in space, typically in a fan shape,¹ has a long history of application to the imaging of fixed structures. Chief among these has been imaging the seafloor to determine the bottom depth over relatively broad transects or swaths. The resulting data are typically georeferenced and displayed as maps, while substantial efforts are made to avoid echoes, also called returns, from scatterers such as fish in the intervening water column.

In other applications, particularly those relating to fisheries, specialized multibeam sonars have been adapted to allow collection of returns from the water column and display of data in real time on a screen, much like conventional sonar data.² The increased areal coverage provided by multibeam sonars without compromise of spatial resolution makes them well suited for studies of fish behavior and vessel avoidance. A number of pioneering studies have looked at the behavior of pelagic fish schools during the passage of an acoustic survey vessel including, for example, Misund and Aglen,³ using the SIMRAD SM600 sonar, with operating frequency of 34 kHz, and Gerlotto *et al.*^{4,5} and Soria *et al.*,⁶ using a RESON SeaBat 6012, with operating frequency of 455 kHz. The behavior of pelagic fish schools during trawling has also been observed with the SM600.³ The swimming speed of fish has been measured using the Doppler function of the same sonar.⁷ Migration of schools of herring (*Clupea*

harengus) in the North Sea has been observed, with concomitant measurements of swimming speed, by both the Simrad SR240 and SA950 sonars, with operating frequencies of 24 and 95 kHz, respectively.⁸ Migration behavior of schools of capelin (*Mallotus villosus*) in the Barents Sea has also been observed with the SR240.⁸ Some other applications of multibeam sonar to fish, as well as limitations in the use of the sonar, have been summarized by Reid.⁹

The three-dimensional nature of multibeam sonar data has also encouraged application of modern visualization techniques, allowing the complex behavior of mid-water targets to be explored interactively and in three dimensions.¹⁰ The advantage of using multibeam sonar for mid-water target identification is also being explored by the military.¹¹

Predator-prey interactions are being studied by means of multibeam sonar. The movements of schools of herring induced by killer whales (*Orcinus orca*) have been measured,¹² as have the interactions between juvenile herring and the Atlantic puffin (*Fratercula arctica*).¹³ In both of these studies, the RESON SeaBat 6012 sonar was used, with respective operating frequencies of 455 and 150 kHz. The behavior of Hawaiian spinner dolphins (*Stenella longirostris*) and pelagic prey has been studied with the new SIMRAD MS2000 sonar, with 200-kHz operating frequency and 128 1.5-deg beams spanning a 150-deg sector.¹⁴

The several fisheries applications mentioned here have succeeded because of the qualitative imaging capability of

the multibeam sonar in a high-signal, low-noise environment, supplemented by registration of the echo magnitude. Calibration is generally important for imaging applications, for example, to define the limits of performance *vis-à-vis* the signal-to-noise ratio and to ensure consistent, distortionless performance at high signal-to-noise ratios. In other desired applications beyond those of visualization,¹⁰ a quantitative capability, hence calibration, is essential. This is true for numerical determination of concentration densities of fish and other biological organisms. It is also true for measurement of bubble populations¹⁵ and particulate concentrations near hydrothermal vents,¹⁶ for example.

There are additional advantages to calibrating multibeam sonars. For example, sonar performance can be monitored, with the ready possibility of detecting changes in this. While a multibeam sonar may be calibrated by the manufacturer prior to shipping to a customer, the owner may want to confirm its compliance with the specifications. Also, the user needs to be able to check the state of the instrument at or near to the time of application.

There are well-established procedures for calibrating sonars. A dozen different basic approaches to measuring the frequency response function of sonar receivers were listed in 1983 by Urick.¹⁷ Given the complexity of multibeam sonar, particularly the large number of beams, use of any of these approaches must be time consuming. At the same time, these calibration procedures require measurements of parts of the sonar, with the inevitable compounding of errors, and use of secondary standards, such as hydrophones, themselves requiring calibration.

Fortunately, there is an alternative procedure available for multibeam sonar, namely that of the standard target.^{18–20} This is widely used for calibrating scientific echo sounders at ultrasonic frequencies, and has already been used for measuring certain performance measures, including the overall system directionality, of the SIMRAD SM2000/200-kHz multibeam sonar^{21,22} and SM2000/90-kHz multibeam sonar.²³

Recognizing the importance of developing simple and straightforward calibration procedures for multibeam sonar, a project has been conducted with the aim of developing convenient calibration protocols for multibeam sonars used for

water-column measurement. Development of new protocols has already been described in a series of preliminary publications.^{21,23,24} This development work is summarized, and the protocols are documented. It is believed that these are sufficiently detailed to serve as guides for the interested user. It is believed further that the protocols collectively will enable multibeam sonars configured for water-column measurement to be used as an integrated set of scientific echo sounders.

In the following, the test multibeam sonars and calibration targets are described, as are the seawater and freshwater facilities and associated instrumentation used in developing the calibration protocols. Methods involving acoustic measurements are reviewed, and results of the development work are summarized in the form of protocols. Some outstanding problems in multibeam sonar calibration are briefly discussed.

II. MULTIBEAM SONARS AND TARGETS

Essential to the development of calibration protocols for multibeam sonar was identification of appropriate sonar systems and specification of standard targets. Both subjects are addressed.

A. Sonars

Three sonar systems were selected for their routine provision of the water-column signal, which is otherwise the exception in most commercial multibeam sonar systems. In this section, details are given on the transducer array geometries. While this information is not usually regarded as being commercially confidential, it is not generally published either, not even in proprietary literature provided to owners of such sonars. The information is valuable in several ways, e.g., for use in beamforming based on individual element signals (Sec. V A) and for other processing or analytical operations, as in calculating nearfield–farfield transition characteristics (Sec. V G).

Some transducer specifications are compared in Table I. Differences in particular configurations are sufficient to require further elaboration in the text.

TABLE I. Comparison of three different multibeam sonars used in the calibration exercises.

Parameter	SM2000/90 kHz	SM2000/200 kHz	RESON Seabat 8101
Operating frequency (kHz)	90	200	240
Receiving array			
No. of channels	80	80	176
Radius (cm)	38.7	20.0	15.6
Angular span (deg)	94.8	88.16	150
Displayed beams			
No. of beams	128	128	101
Angular resolution (deg)	1.5	1.5	1.5
Displayed-beam signals			
No. of bits	8	8	8
Dynamic range (dB)	48	48	48
Individual-channel signals			
No. of bits	12	12	8
Dynamic range (dB)	72	72	48

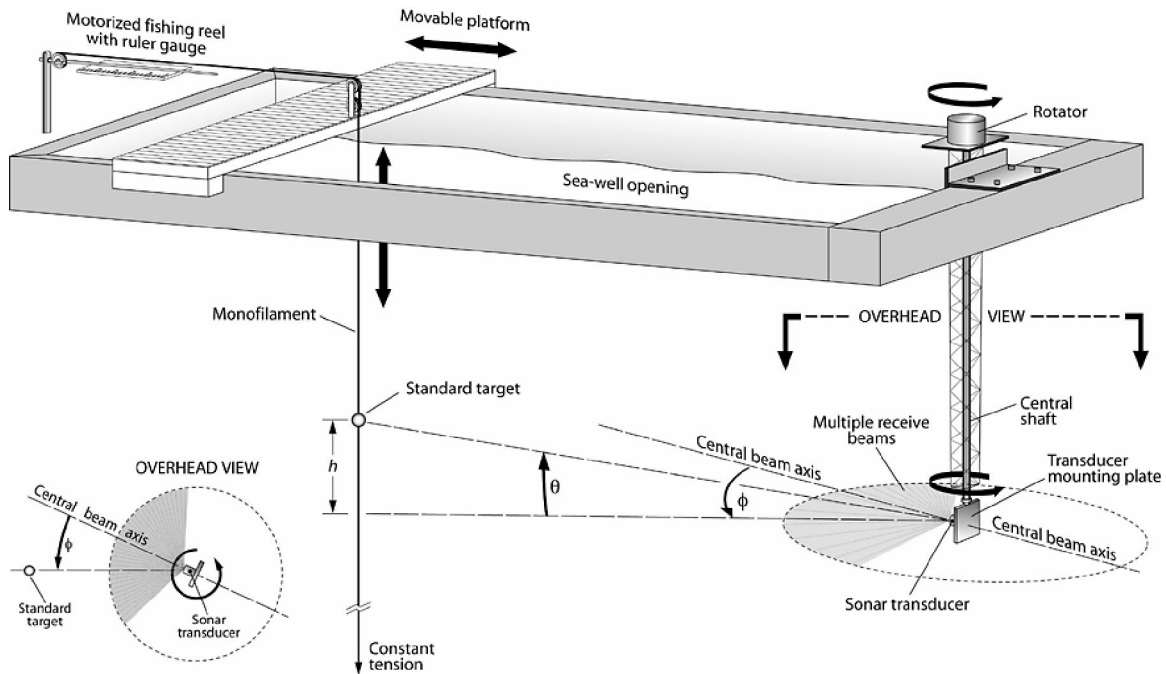


FIG. 1. Schematic diagram of the sea well as configured for calibration. For the transducer mountings, as indicated in Fig. 2, the equatorial plane is that plane running through the arc of the receiving elements, with $\theta=0$ or equivalently $h=0$, thus coinciding with the horizontal plane. The polar planes are perpendicular to the equatorial plane and contain the rotational axis of the transducer; for the particular configuration, they are vertical planes.

Transducer-mounting convention: In all cases, the transducers were mounted with their active surfaces in the vertical plane, as indicated in Fig. 1. Thus, the acoustic axis of each was horizontally oriented.

Terminology: As indicated below, the elements of each receiving array used in this study are arranged in a single row along a circular arc. The plane of this arc is called the equatorial plane. The perpendicular planes containing the rotational axis of the transducer are called the polar planes. Each particular polar plane is distinguished by its azimuthal orientation, which may define an acoustic axis if the beam-forming is done through postprocessing.

1. SIMRAD SM2000/200-kHz multibeam sonar

The SM2000 has two separate operating modes, referred to as imaging and echo-sounding. In the imaging mode, the same transducer array is used in both signal transmission and echo reception. This array consists of 80 elements positioned with equal spacing on a circular arc (Fig. 2), thus defining the equatorial plane. Each element is rectangular, 2.54 by 19.05 mm. Two versions are fabricated. For that of nominal span 120 deg, the radius to the outer faces of the elements is 20.00 cm, and the interelement spacing is 1.12 deg, hence spanning a total angular sector of 88.16 deg as measured between the centers of the end elements. For that with nominal span 150 deg, the radius to the outer faces of the elements is 10.85 cm, and the interelement spacing is 1.13 deg, spanning a total sector of 155 deg. In the associated digital signal processor, a total of 128 receive beams are formed, each with beamwidth 1.5×20 deg in the respective horizontal and vertical planes.

In the echo-sounding mode, the described transducer array is used only in reception. Transmission is effected by an

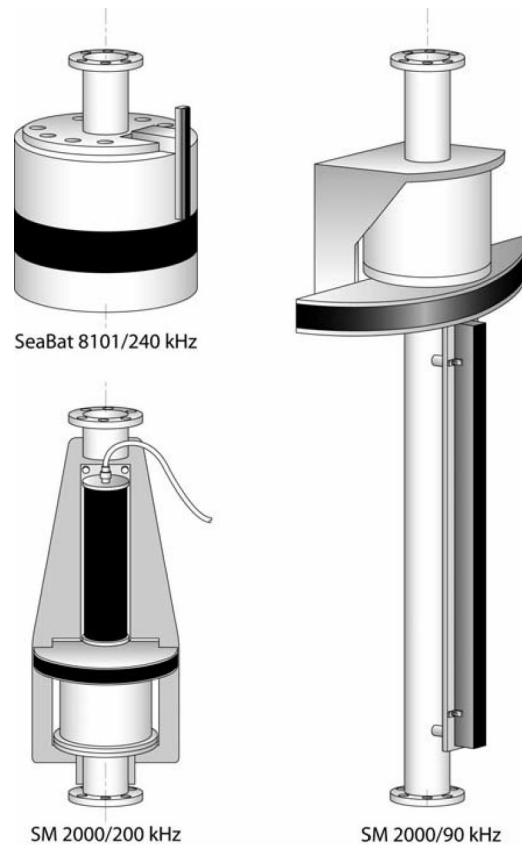


FIG. 2. Schematic diagrams of transducer-mounting plates for the SeaBat 8101 240-kHz multibeam sonar, SM2000/200-kHz multibeam sonar, and SM2000/90-kHz multibeam sonar. The black rectangular areas designate the protective coverings of the external transmitter. The black curved areas designate the same for the receiving arrays. In the case of the SM2000 models in the imaging mode, the external transmitter is not used, and the curved arrays are used both for transmission and reception. In the case of the SM2000 models in the echo-sounding mode, the external transmitter is used.

external line array of 50 elements of total length 27.6 cm oriented along the vertical, thus in the central polar plane perpendicular to the equatorial plane of the receiving array (Fig. 2), and thence forming a Mills cross.¹⁷ Each transmitting element is a narrow band of width 2.1 mm spanning 177 deg of a circular arc of radius 44.45 mm. The entire array appears to be a half-cylinder. The elements are stacked with a center-to-center distance of 5.6 mm. They are weighted in amplitude, i.e., amplitude shaded.¹⁷ The computed nearfield distance is 11 m.²⁵ Through digital signal processing, 128 receive beams are formed for the display, with nominal beamwidths 1.5×1.5 deg. The number of beams that can be formed on an external processor is arbitrary.

The distinction between the two modes is recognized by a particular naming convention, which is adhered to in this work. For the general system, with optional use of either mode, reference is made to the SM2000 multibeam sonar. For the particular configuration in which the external transmitter is used, reference is made to the SM2000 multibeam echo sounder.

The transmitted signal in both modes is typically a finite, or pulsed, sinusoid at 200 kHz. The duration may be specified directly by the user or indirectly based on the specified range. For target ranges of order 5–25 m, which are typical in standard-target calibrations of high-frequency sonars, a pulse duration of 100 μ s was generally used. The pulse repetition frequency was determined automatically in accordance with the range, typically about 1 Hz or 1 ping/s. Notwithstanding the appearance of user control over the settings, the sonar itself is configured internally with certain constraints that cannot be exceeded. In addition, the sonar can neither accept nor generate an external trigger signal, which may be regarded as highly inconvenient in many applications.

The sonar displays echoes as received on 128 beams over the total angular sector of 120 or 150 deg. Significantly, the sonar can record either the raw complex echo signals from each of the 80 elements or the 128 beamformed signals. Beamforming performed within the sonar digital signal processor (DSP) uses a Hamming-type amplitude-weighting.²⁶ Beamforming performed through postprocessing external to the sonar DSP uses Chebyshev-type amplitude-weighting.¹⁷ Other shading schemes may be implemented when using the individual element outputs, and other apertures may also be defined.

The data are digitized with 8-bit resolution for display and with 12-bit resolution for recording purposes. The nominal dynamic range of the recorded data is thus 72 dB.

2. SIMRAD SM2000/90-kHz multibeam sonar

The architecture of this sonar, which operates at 90 kHz, is similar to that of the 200-kHz sonar described above. Similarly to the 200-kHz sonar, the 90-kHz sonar operates in both imaging and echo-sounding modes. The collocated imaging array (Fig. 2) also consists of 80 identical elements spaced at intervals of 1.20 deg along a circular arc of radius 38.74 cm measured to the outer face of the element and

spanning 94.8 deg as measured between the centers of the end elements. The dimensions of the individual rectangular elements are 5.08×64.8 mm.

The external transmitter is a linear array 86.6 cm in length oriented perpendicularly to the equatorial plane of the receiving array (Fig. 2). This consists of 12 pairs of elements. Each element is rectangular with dimensions 3.8×64.8 mm. The center-to-center distance between elements in a pair is 17.8 mm. The center-to-center distance between pairs, aligned end-on, is 65.3 mm.

The digital signal processor forms 128 beams for display. The nominal beamwidths are 1.5×20 deg in the imaging mode, without use of the external transmitter, and 1.5×1.5 deg in the echo-sounding mode, with use of the external transmitter.

Transmit signals are pulsed sinusoids with a center frequency of 90 kHz. For nominal target ranges of order 5–25 m, which are used in standard-target calibrations, a pulse duration of 300 μ s was typically selected.

As with the 200-kHz sonar, the 90-kHz sonar provides both the beamformed signals and individual element signals, but only one at a time depending on operator choice. Thus, the user can choose to perform the beamforming independently of the sonar processor.

As with the 200-kHz multibeam sonar, the digitized data are displayed with 8-bit resolution, but recorded with 12-bit resolution. Thus, the nominal dynamic range of the recorded data is 72 dB.

3. RESON SeaBat 8101 multibeam sonar at 240 kHz

This sonar consists of separate transmitting and receiving transducer arrays (Fig. 2). The transmitting array is a linear array of 36 elements conforming to a slender cylinder of radius 2.5 mm, each 5.28 mm in length and separated by a gap of 0.7 mm, hence with total length 21.46 cm. The transmitting array is oriented perpendicularly to the equatorial plane of the receiving array. A swath exceeding 150 deg is insonified. Shading is employed to reduce the on-axis transmit signal strength, as this typically corresponds to the nadir in bathymetric mapping exercises. The same shading achieves a one-way, received sidelobe level of about -20 dB, as computed.

The receiving array is composed of 176 elements evenly spaced on a circular arc of radius 15.6 cm. Each receiving element is rectangular, with height 16.88 mm and width 3.1 mm. The center-to-center distance subtends an arc of 1.5 deg. The system forms 101 1.5×1.5 -deg beams over an angular sector of 150 deg.

In contrast to the SM2000, the SeaBat 8101 operates in only one mode, with separate transmitting and receiving arrays. Thus, it functions like the SM2000 multibeam echo sounder, but is referred to in this work as a multibeam sonar, consistent with the manufacturer's own naming convention.

Although the SeaBat 8101 does not support external-trigger operation, it does accept a synchronized clock signal from another computer. The computer that transmits the synchronization signal can also accept motor-position information from the motor controller. The beamformed data are

digitized and stored with 8-bit resolution. The nominal dynamic range is thus 48 dB.

B. Standard targets

A number of standard targets were selected or devised. In every case these were based on the spherical form. Principles of optimization of the diameter for both narrow-band and broadband sonars^{18,27–29} were not applied in the present work. Rather, for reasons of convenience, a number of sub-optimal targets were chosen based on ready availability. These included spheres made of electrolytic-grade copper with diameters 23 and 60 mm, tungsten carbide with 6% cobalt binder with diameters 20 and 38.1 mm, aluminum with diameter 60 mm, and titanium with diameter 47 mm.

To achieve substantially higher target strengths, several focusing spheres^{30–33} were fabricated.³⁴ Two of these were based on spherical shells made of high-silicon bronze with external diameters 101.6 mm and shell thickness 3 mm. One was outfitted with a movable piston plug, diameter 14.22 mm, and the other with a simple shallow screw. Another focusing sphere was fabricated from a hollow rubber ball of diameter 88.9 mm, with thickness varying from 1.7 to 2.1 mm (average 1.9 mm). The three hollow spheres were vacuum-filled with the fluorocarbon liquid FS-5, also known by the tradename Fluorlube.³⁵

In some cases, the target strengths of the several spheres were either computed according to the respective theories for solid elastic spheres,^{36,37} noting corrections,¹⁸ or the appropriate limiting case as given by Goodman and Stern,³⁸ or for spherical shells,³⁸ assuming accepted values for the material properties. In other cases, the target strengths were determined according to comparative measurements. The results are shown in Table II.

III. CALIBRATION FACILITIES AND THEIR INSTRUMENTATION

Both seawater and freshwater facilities were instrumented for the work of developing calibration protocols. Earlier preliminary reports^{34,39,40} are summarized and updated.

A. Sea well

Iselin Dock at the Woods Hole Oceanographic Institution is a large dock supported by concrete pilings. Large reinforced concrete blocks have been removed from a particular area, exposing a sea well with rectangular opening $6 \times 13 \text{ m}^2$ (Fig. 1). The hard bottom beneath this opening slopes from about 15 to 20 m. Another area approximately 22–24 m from the transducer mounting apparatus is also easily accessible for mounting equipment such as calibration targets.

The sea well is within about 700 m of the Woods Hole passage. Because of its high-tidal-current velocities,⁴¹ stratification is at most weak. Because of the performance of sea-well calibrations over quite short ranges, effects of such stratification are negligible.

Interestingly, the tidal current at Iselin Dock is rectified. The flood occurs with substantial current towards the south-east, typically of the order of 1.5 knots. During the ebb, however, the water level falls or settles with very slight and variable current. The tidal range is typically of order 0.8 m. It has been possible to perform measurements during both tidal states, with those made during the flood being less variable.

The sea well was equipped with a variety of instruments and gear for deploying and rotating sonar transducers and for suspending and controlling targets at known positions in the transducer beam. Because of the wide range of uses of the sea well, all mountings were made in a provisional manner, but with drilled holes or markings for future reassembly of the facility.

A standard, 6-m-long shipping container was provided or configured as a laboratory for the electronic instruments and computers during each calibration trial. Mains power of both 110 and 220 V at 60 Hz was supplied by cable. The laboratory was placed immediately beside the sea well for operational convenience, specifically to minimize the cable distance to the transducers and rotation apparatus and to facilitate communication among the various participants.

1. Transducer-mounting system

During calibration trials, a 6-m-long steel-trussed tapered antenna tower weighing 150 kg was affixed to a mas-

TABLE II. Nominal target strength (TS) values of standard targets as computed and as inferred at 200 kHz by means of the SM2000 multibeam echo sounder. The TS values shown in the parentheses have been computed at precisely 200 kHz according to the single-frequency theory³⁶ based on tabulated values for the physical properties of the respective target and applicable conditions of temperature and salinity. The inferred TS values have been calculated relative to the average target strength of the 38.1-mm tungsten carbide sphere (WC38.1).

Target	Diameter (mm)	Relative TS (dB)	Inferred TS (dB)	Computed TS (dB)
AL60	60	96.52 ± 1.83	−35.31	−34.46
WC38.1	38.1	96.41 ± 1.02	−35.42	−34.46
		92.43 ± 0.76	−39.38	−39.36
		92.47 ± 0.87	−39.34	−39.36
WC20	20.0	88.99 ± 0.23	−42.84	−44.99
High-silicon bronze	25.4	94.87 ± 0.17	−36.96	
Focusing sphere 1	101.6	96.74 ± 0.95	−35.09	
Focusing sphere 2	88.9	93.47 ± 1.28	−38.36	
Stainless steel	50.8	93.40 ± 1.47	−38.43	
Titanium	47	84.88 ± 0.60	−46.95	

sive I-beam spanning the sea-well opening at the northwest end. The tower was held securely in place by a mechanical lock under gravity. The tower was strapped to the lower part of the I-beam mounting plate to resist the bending moment caused by the flooding tide when acting on the tower and transducer submerged at nominal 3-m depth.

An aluminum pipe is centrally mounted in the tower to serve as a shaft for holding and rotating the transducer. The shaft is guided by holes in the steel endplates and an intermediate plate added to reduce flexure. Low-friction bearings are mounted at each hole.

The shaft is held by the rotation apparatus, described in the next section. A flange with six bolts is mounted on the lower end to serve as a universal coupling joint for the transducer plate.

Given significant differences in the shape and size of each of the multibeam sonar transducers, a separate mounting plate was fabricated in aluminum to hold the respective transducer in a secure manner. Longer mounting plates for holding separate transmitting and receiving arrays were generally braced with aluminum angle-brackets for stiffness under flooding conditions. Examples of mounting plates with attached transducers are shown in Fig. 2.

In the case of separate transmitting and receiving transducer arrays, alignment of the several beams could be effected by shimming one of the arrays relative to the other. This was done for the SM2000/90-kHz multibeam echo sounder to reduce effects of parallax given the relatively short range of the target.

2. Rotation apparatus

A commercial antenna motor made by M2-Antenna Systems, the MT1000 antenna azimuth positioner, was adapted to serve as a precision rotator. The apparatus was mounted atop the central shaft of the tower, supporting both thrust and radial loads.

In the course of adapting the rotator, the central gear was replaced by one made of marine bronze. The original dc-motor was also replaced with an ac-motor with 14-bit absolute encoder and counter. The rotator can turn about a 360-deg arc, thus achieving a nominal precision of ± 0.01 deg, well within the design goal of ± 0.05 deg.

The accuracy of the rotator was established by measuring the deflection of the beam of a laser pointer mounted atop the central transducer-rotation shaft. Under commanded rotations, the position of the center of the red laser spot on the bridge platform, for example, was determined to within ± 5 mm at a nominal range of 11.7 m. This was found to be repeatable over a period of days, hence with demonstrable accuracy to within ± 0.025 deg.

The alignment of the transducer at the end of the 6.5-m shaft with the rotator was confirmed by attaching an outrigger to the rotator and dropping a plumb line from the outrigger to a corresponding outrigger affixed to the shaft immediately above the transducer plate, while the transducer tower was held vertically by a mobile crane. The distance from the shaft centerline to the plumb bob centerline was 47.1 cm. Very small adjustments could be made to the alignment by exploiting the small degree of play in the flange coupling. It

is estimated that the transducer could be aligned to within ± 0.5 mm over 4 m, or to within ± 0.007 deg.

Rotator positions were sent in ASCII format via serial link, RS232, to the SM2000 and were recorded with the SM2000 raw data files. There was a five-ping delay between the actual rotator position and the ping that recorded the position.

3. Target-suspension and -positioning systems

A number of systems were procured or assembled for suspending standard targets and controlling their positions in the transducer beam. One system was adapted from that currently used on board fisheries research vessels.^{19,20} The target was held by each of three monofilament nylon lines connected to motorized fishing reels, with outriggers mounted on barrier timbers, two on one side of the sea well and one on the opposite side. The position of the target was controlled by means of joysticks from the electronics laboratory.

In a second system, indicated schematically in Fig. 1, the target was suspended by a central monofilament line spooled on a motorized fishing reel. A second monofilament nylon line was attached to the central line and threaded through a block resting on the bottom of the sea well directly below the target to a second block. This was attached to a safety beam bordering the well, with lead weight suspended at the submerged end of the line to maintain constant tension, whatever the tidal state.

A third system was derived from the second by attaching four additional lines to the target.³⁴ Each of these was tensioned by running the respective line through a block to an attached weight. Each block was supported by the sea well bulkhead mounted along the long sides. The four blocks were arranged at the corners of a square. This additional tensioning was obliquely upwards, offering stability against lateral, tidally induced currents.

The second system was most used in the trials. The vertical suspension line was run through a block mounted on a platform bridging the sea well to a motorized fishing reel (Fig. 1). A variable voltage motor was used to control the fishing reel. The line was run directly over a meter ruler, allowing depth adjustments to within ± 0.5 mm. At a range of 11.7 m, this corresponds to an angular uncertainty of ± 0.003 deg.

B. Freshwater tank facility

The Chase Ocean Engineering Laboratory at the University of New Hampshire (UNH) has a freshwater tank that is nominally 12 by 18 m in plan view, with 6-m water depth. Modifications were made to this tank to facilitate the required acoustic measurements. These modifications fundamentally concerned the placement and control of the transducer under test in the tank. There were three fundamental issues to consider in achieving this goal: *X-Y-Z* transducer-positioning, rotation, and system integration.

1. *X-Y-Z* transducer-positioning system

The *X-Y* positioning in the plan view of the tank was accomplished through the addition of an industrial crane-rail system spanning the 12-m width of the tank, and moving

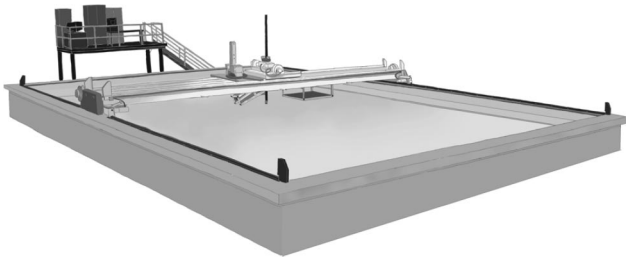


FIG. 3. Schematic diagram of the freshwater tank as configured for calibration.

along the longer, 18-m dimension (Fig. 3). The span accommodated a carriage, which moved across the 12-m span and was designed to secure the rotator. The rotator was mounted on the carriage and held the transducer under test, while controlling the incremental rotation about the vertical or Z axis. The rotator held a carbon fiber pole to which the transducer under test was attached (Fig. 4). The Z depth of the transducer was adjusted by securing the pole in the three-jaw chuck at the appropriate location.

2. Rotator

Rotation was accomplished using a Yuasa International, three-jaw chuck rotator (SUDX-320) with a programmable controller (UNDC-100). This rotator was selected due to the zero-backlash claim, and the ability to effect small increments of angular rotation, 0.1 deg. The rotator was readily mounted on the carriage. The transducer-mounting pole and rotation shaft was placed in the three-jaw chuck. A universal transducer-mounting flange was attached to the bottom of the pole. The pole was fabricated from carbon fiber, and the length was adjusted by adding sections outfitted with locally fabricated centering couplings.

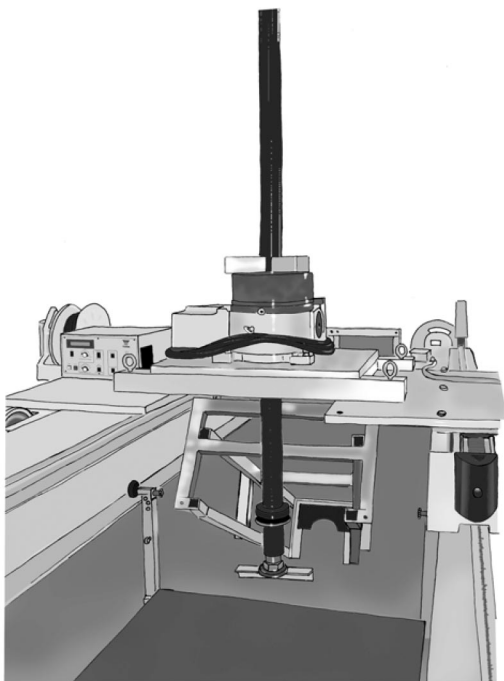


FIG. 4. Configuration of the transducer and its mounting system in the freshwater tank.

The rotator performance was verified by means of a machinist dial indicator to measure the arc length traveled due to an incremental angular input. Two-degree arcs were randomly selected around the circumference. Each 2-deg arc was traversed with 0.1-deg increments in both directions. Precision control over the angle and the ability of the rotator to return to the place where it began, to within the accuracy limit of ± 0.1 deg, were both demonstrated. Details of this system development are reported in Baldwin *et al.*⁴⁰

3. System integration

System integration was achieved using LabView© software. An interpreter program was acquired which enabled LabView to communicate with the rotator controller. Subsequently other code was developed to facilitate the tasks required to perform a beam pattern measurement. Transmit signal generation, receive signal acquisition, transmit-receive-rotation coordination, and beam-pattern plotting were the basic functions. The specific linking of these functions was an artifact of the particular source-receiver-target or hydrophone configuration.

4. Target suspension and positioning

Target suspension and positioning were effected with a single line of monofilament nylon deployed through a block at the end of an outrigger. The line was attached to a motorized fishing reel, allowing fine control of the target depth, as described in more detail above in Sec. III A 3.

IV. METHODS

A number of procedures preparatory to calibrating sonars have already been described in the previous section. These include alignment of the transducer and rotator and possible shimming of separate transmitting and receiving arrays.

Another general preparatory measure is that of measuring the hydrographic state of the calibration environment. Vertical temperature and salinity profiles should be taken at intervals. Additionally, time series of temperature and salinity might be recorded at the transducer depth. During the particular calibration trials in the sea well, vertical conductivity-temperature-depth (CTD) profiles were taken, indicating that the seawater was well mixed. At other times, the CTD sensor was placed at the depth of the transducer to collect time series of temperature and salinity. The average temperature, salinity, and sound speed measured throughout one trial were 2.88 ± 0.02 °C, 31.60 ± 0.25 PSU, and 1457.5 ± 0.4 m/s, respectively.

Methods involving acoustic measurements are described in this section. Some of these are common to calibration of both scientific echo sounders and multibeam sonars; they are included here for the sake of completeness.

Implicit in the measurements for the SM2000 multibeam echo sounders are real-time storage and rapid retrieval of the echo signals from the individual receiving transducer elements and availability of operational software for beamform-

ing. Algorithms for beamforming have been described in some detail for multibeam sonar.^{21,22} These follow the general method.^{17,42}

A. Target positioning and mapping central polar beam pattern

A standard target can be placed at an essentially arbitrary position in the sonar beam by purely geometrical considerations. Not all positions are equally favorable, it must be emphasized. It is generally desirable to avoid reverberation from surfaces and possible extraneous scatterers in the water column, e.g., fish and underwater moorings. It is also desirable, if not absolutely necessary, to make measurements in the transducer farfield.

In practice it is easier to perform relative positioning of the target in the transducer beam by placing the target at the approximate depth of the centered transducer, then finding the coincident horizontal and equatorial planes of maximum sensitivity by moving the standard target systematically through a series of depths and making multiple measurements at each target depth. Beamforming is then performed, and the resulting data are fit in a least-squares sense by a quadratic or other nonlinear function. The beamformed data describe the central polar beam pattern.

Examples of resulting data derived from measurements with the SM2000/90-kHz multibeam echo sounder are shown in Fig. 5. In the first example, taken from the sea-well facility, a standard target at 23-m range was lowered, and then raised, systematically at constant increments of 10 cm or 0.25 deg. In the second example, taken from the freshwater-tank facility, the same standard target at 11.7-m range was used and the procedure was repeated, but at 2-cm increments, hence with nominal resolution of 0.10 deg.

B. Measurement of system response and choosing receiver gain function

The performance of a system is considered to be linear if a change in the amplitude of an input signal results in a linearly proportional change in the amplitude of the output signal. While there is no need for a system to perform linearly for it to be used quantitatively, given that it performs monotonically, the system response needs to be determined. Among other reasons, the sonar will be most useful if its gain function enables both weak and strong echo signals from interesting targets to be registered without loss due to inadequate sensitivity or saturation, with clipping, due to inadequate dynamic range. The general effect of clipping may be *hard*, with abrupt change in any amplitude exceeding a threshold value to the threshold value itself, or *soft*, with gradual approach to the threshold value, where the exact relationship is nonlinear. The general case has been treated under the name of band-pass limiting.⁴³ The importance of choosing the gain function is evident from practical considerations of dynamic range: the test multibeam sonars used in this study have rather limited dynamic ranges, especially relative to those of modern scientific echo sounders.⁴⁴

Multibeam sonars generally have a number of available transmit-power settings. In keeping with the aim of develop-

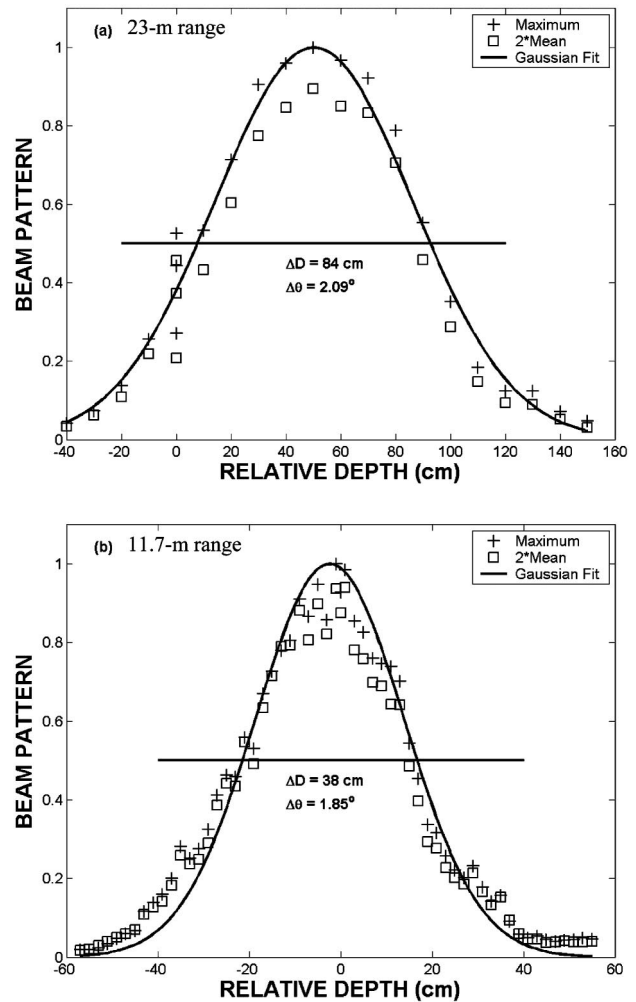


FIG. 5. Directionality of the SM2000/90-kHz multibeam echo sounder in the vertical plane, as determined by changing the depth of the rubber-walled focusing sphere (a) at 23-m range in 4-cm increments in the sea well and (b) at 11.7-m range in 2-cm increments in the freshwater tank. The directionality is expressed through the beam pattern in normalized intensity units. The maximum intensity value is that observed in a spatial window consisting of five contiguous beams and five range cells centered on the respective beam and range of nominal greatest sensitivity. The mean is that of all echo values within the window. A Gaussian function has been fitted to the maximum intensity values.

ing convenient and useful protocols, the transmit power of each sonar was set at that level most often used in field applications.

Measurement of the system response is performed by observing the change in echo amplitude due to changing the receiver gain function of the sonar. It is effected conveniently by observing the change in echo amplitude from a standard target under systematic changes in the receiver gain function. This is illustrated for the SM2000/90-kHz multibeam echo sounder in Fig. 6, where the full range of gain settings has been exercised. A linear function has been fit to the echo-intensity data by means of a least-squares-regression analysis.⁴⁵ On the basis of the data and their analysis, the gain function $20 \log r + 2\alpha r$ was chosen, where r is the range and $\alpha = 0.0157 \text{ dB/m}$. In the case of the SM2000/90-kHz multibeam echo sounder, the system sensitivity was expressed through a numerical value. This was determined from the equation

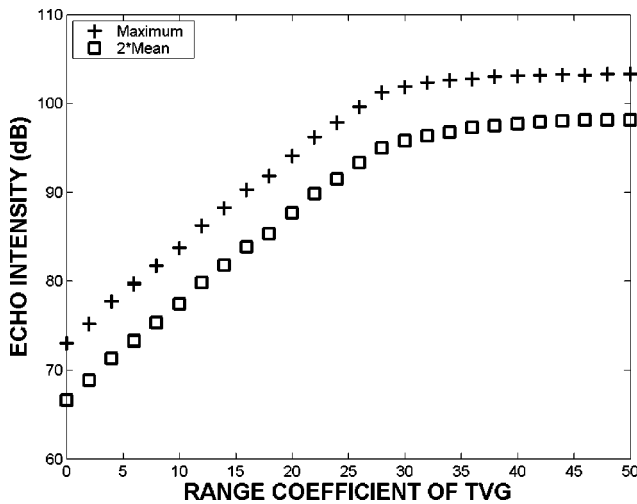


FIG. 6. Measured dependence of the echo intensity from the standard target AL60 on the gain function of the SM2000/90-kHz multibeam echo sounder. The target was measured at 11.7-m range in the sea well, with pulse duration of 300 μ s. For an explanation of maximum and mean values, see the caption to Fig. 5.

$$TS = 10 \log_{10}(G(r_0)|V_{rec}|) + 20 \log_{10}(r_0) + 2\alpha r_0, \quad (1)$$

where $TS = -34.40$ dB denotes the target strength of the 60-mm-diam aluminum sphere (Sec. II B) at 90 kHz for the particular hydrographic conditions, $r_0 = 11.7$ m, V_{rec} is a relative digitized real number that is proportional to the received voltage level, and $\alpha = 0.0157$ dB/m. It is emphasized that V_{rec} also depends on the instrument settings, including transmit power level, pulse duration, sampling rate, and assumed absorption coefficient value, which in the present case was 0.0068 dB/m. Substituting for V_{rec} and solving, the sensitivity coefficient is $G(r_0) = 3.68 \times 10^{-6}$. Multiplying the raw echo amplitude due to an arbitrary target or targets at the same 11.7-m range, with the same gain function, yields backscattering units of square meters. For other gain functions at the same range, the relationship defined in Fig. 6 may be used. Since the calibration was performed in the transducer nearfield, Eq. (1) must be modified for use at other ranges.

C. Measurement of system stability

There is little reason to suspect instabilities in performance of multibeam sonars designed for scientific use, although their possible occurrence at high pulse repetition frequencies might be entertained. Ultimately, it is only measurement that can decide the matter or quantify the degree of stability.

Measurement of the system stability is straightforward. The target is held at a fixed position in the transducer beam, and echo measurements are performed over a relatively long period of time. This might last from minutes to hours or days or weeks, the longer period corresponding to potential use on fisheries survey cruises.

An example of a measurement of system stability is presented in Fig. 7. This shows the time series of echo amplitude from the SM2000/200-kHz multibeam echo sounder due to repeated insonification of the standard target WC38.1.

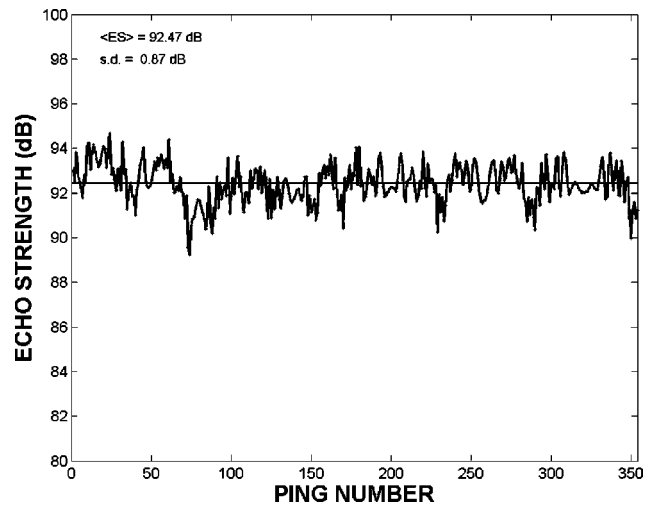


FIG. 7. Echo amplitudes from an on-axis standard target, WC38.1, as measured at 11.7-m range in the freshwater tank by the SM2000/200-kHz multibeam echo sounder over a time period of 175 s.

D. Measurement of equatorial-plane sensitivity

For conventional imaging purposes, the equatorial-plane sensitivity function is very important. The value of this function for a small target at a given angle is the peak echo amplitude from among all beams where the target is detected. It is advantageous in imaging applications that the function be relatively flat over the full angular span of the sonar.

The equatorial-plane sensitivity function must vary over the angular span, for it is defined by a finite series of beams of fixed orientation. Adjacent beams meet at a level that depends on the number of beams, their total angular span, the respective aperture size, and the shading of individual elements. For the 128 displayed beams of the SM2000, this level is roughly 2 dB less than peak levels, thus defining a measure of ripple.

For scientific imaging purposes, knowing the sensitivity function and how it varies is important, for it enables compensation to be effected and distortions removed.

The equatorial-plane sensitivity function can be determined by measuring the echo response of the sonar at closely separated, discrete angles over the full angular span. This was done for the SM2000 multibeam echo sounders and the RESON SeaBat 8101 multibeam sonar. Exemplary resulting functions are shown in Fig. 8.

E. Measurement of central-region directionality in two dimensions

As mentioned in the previous section, there is particular interest in the equatorial-plane sensitivity function. This is a one-dimensional slice of the more general two-dimensional sensitivity function. Implicit in this description is a rather broad width of the beam in the transverse direction, nominally 20 deg in the vertical plane for the imaging mode of the SM2000. In the echo-sounding mode, the nominal beamwidth is 1.5 deg in the vertical plane.

Again, there is no substitute for measurement of the directionality. This has been done in two dimensions by repeat-

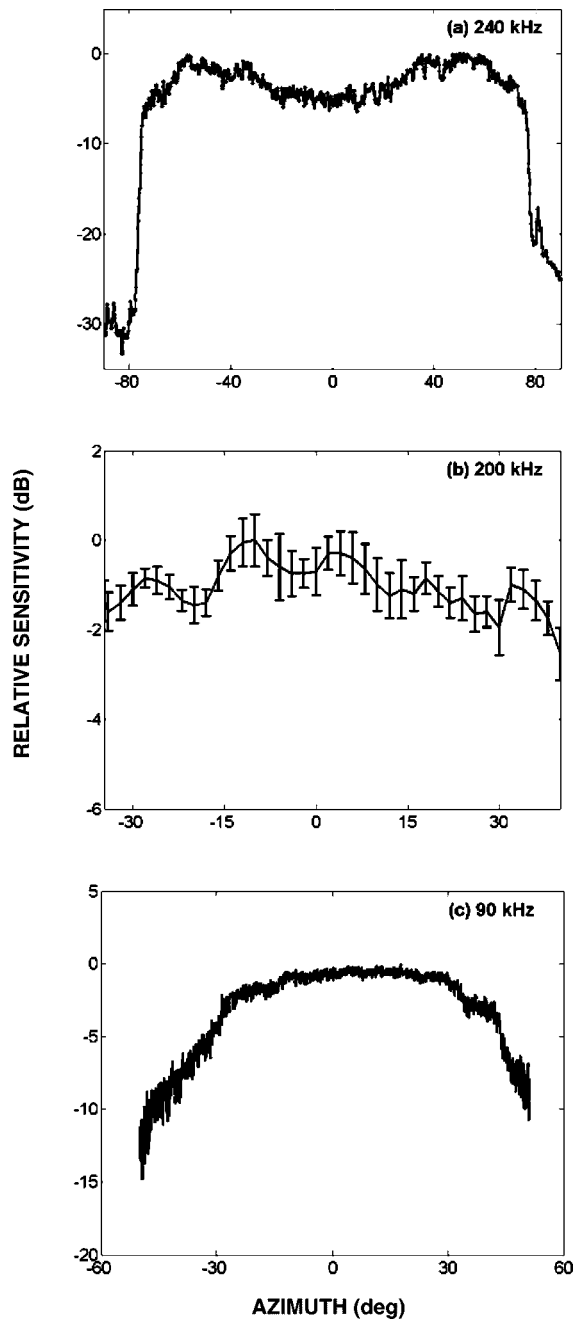


FIG. 8. Measured equatorial-plane sensitivity function of the (a) Seabat 8101 240-kHz multibeam sonar, (b) SM2000/200-kHz multibeam echo sounder, and (c) SM2000/90-kHz multibeam echo sounder. The target range was 11.7 m. In (b), only part of the central section of greatest sensitivity was measured.

ing the one-dimensional measurements, with the standard target moved systematically through a series of discrete vertical positions.

Exemplary results of this measurement for the SM2000/90-kHz multibeam echo sounder are shown in Fig. 9. Measurements are shown of the central region of directionality.

F. Measurement of directionality of individual beams

An envisioned major use of multibeam sonar in fisheries survey work is quantification of fish density. For measurements of resolved single fish by echo counting, this requires

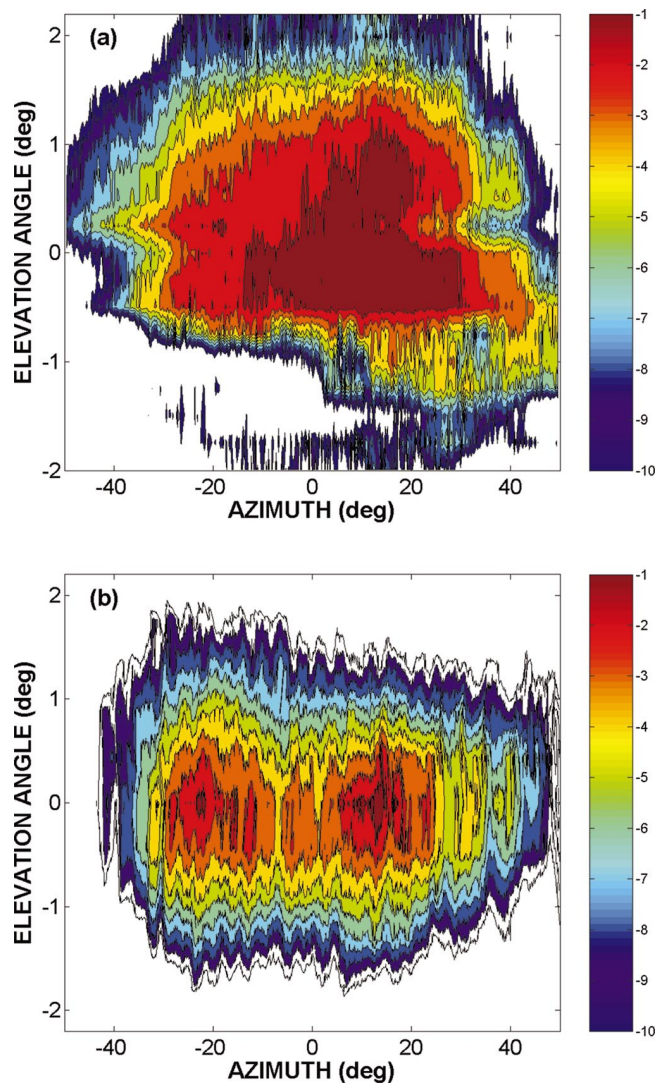


FIG. 9. Measured central-region directionality in two dimensions of the SM2000/90-kHz multibeam echo sounder at (a) 23-m nominal range in the sea well and (b) 11.7-m nominal range in the freshwater tank. The dark brown color represents echo levels extending from -1 to 0 dB.

knowledge of the acoustic sampling volume,⁴⁶ which depends on the individual beam patterns, i.e., the beam patterns of individual apertures used in the quantification. For measurements of unresolved fish aggregations by means of echo integration,¹ measurement of the area backscattering coefficient requires knowledge of the equivalent beam angle. This is the integral of the product of transmit and receive beam patterns, each expressed in the intensity domain, over all angles.¹⁷ For directional beams, this angular measure can be determined with high accuracy from the characteristics of the main lobe alone.

Data collected in the measurement of the central-region directionality in two dimensions, described in the previous section, are sufficient to determine this function. Individual-aperture beam patterns can also be derived by approximation from measurement of the equatorial-plane sensitivity function, and assumption that the beamwidth in the vertical plane is that measured in the central polar beam pattern (Sec. IV A). In either case, the individual beam patterns are formed

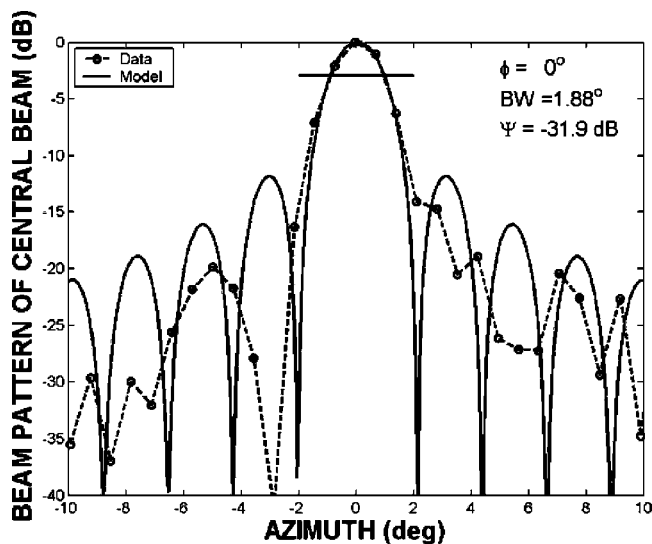


FIG. 10. Equatorial-plane beam patterns, measured and estimated, of a single aperture of the SM2000/90-kHz multibeam echo sounder. The measured value of the beamwidth BW and the estimated value of the equivalent beam angle Ψ are attached to the respective function.

without further selection or reduction, as in the operation of selecting peak values.

Results from the second procedure are shown in Fig. 10 for a single aperture and in Fig. 11 for a set of apertures. Individual beam patterns are shown in the horizontal, equatorial plane together with fitted functions. The inferred equivalent beam angle for the central beam in Fig. 10 is $7.76 \cdot 10^{-4}$ sr or, logarithmically, -31.1 dB. The theoretical prediction based solely on the geometry of the transducer array is $6.46 \cdot 10^{-4}$ sr or -31.9 dB.

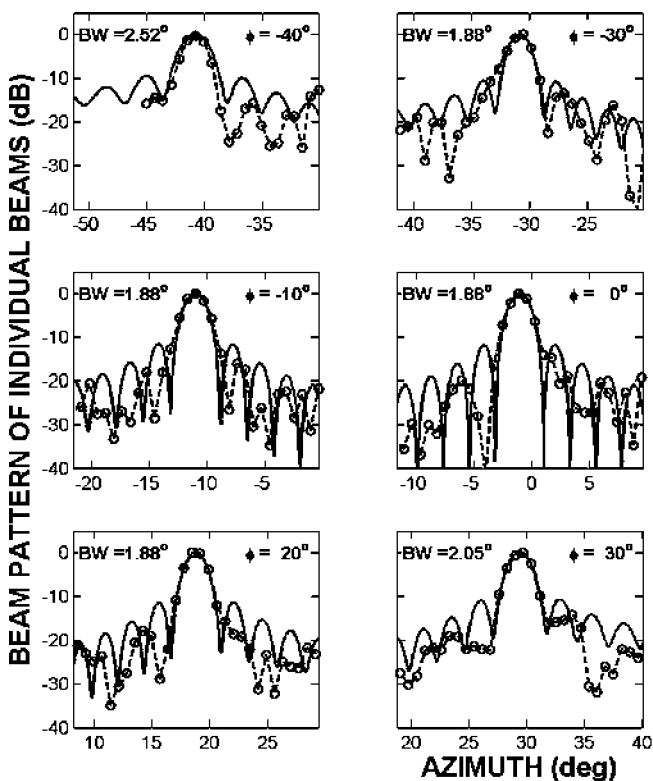


FIG. 11. Equatorial-plane beam patterns, measured and estimated, of a set of individual apertures of the SM2000/90-kHz multibeam echo sounder. Measured values of the beamwidth BW are shown.

G. Measurement of system axial sensitivity in transducer nearfield

It is often advantageous to make measurements of beam patterns in the transducer farfield. At the so-called farfield distance, the radiated field in terms of pressure or particle velocity amplitude decreases inversely as the range r . This farfield distance is conventionally defined as one-half of the square of the maximum transducer dimension divided by the wavelength.²⁵ When the maximum measurement range available for calibration is less than the farfield distance, it may be convenient or necessary to make measurements in the transducer nearfield. Such measurements can be productive if the transducer nearfield characteristics can be modeled successfully, then extrapolated to the farfield.

An illustration is provided by calibration of the SM2000/90-kHz multibeam echo sounder with external transmitting transducer in the freshwater tank. Given the standard array length, 86.6 cm, the farfield distance is about 23 m, which can be achieved without significant interference in the sea well but not in the freshwater tank. Measurements were therefore performed of the system axial response in the transducer nearfield.²³ These are documented in Fig. 12. The modeled results are superimposed, suggesting the possibility of making angular measurements at nearfield ranges as close as 6 m, then extrapolating these by means of modeling to the ordinary farfield.

V. PROTOCOLS

Protocols are developed for each of the several measurement methods, with the intention of describing how to perform these in the form of prescriptions. The protocols are preceded by a description of preparatory procedures to be

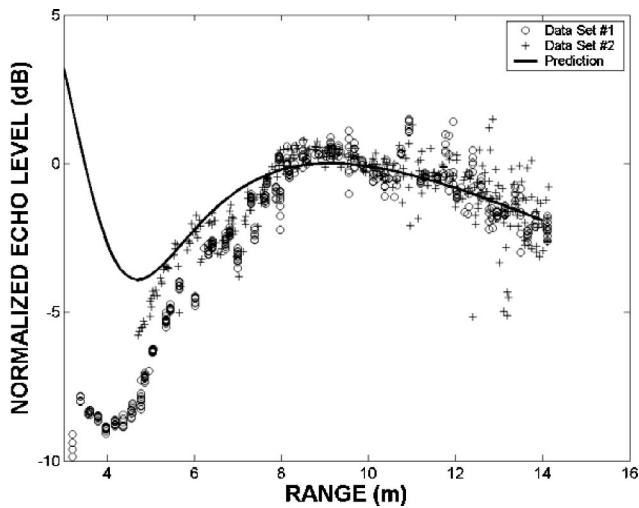


FIG. 12. Dependence of the axial response of the SM2000/90-kHz multi-beam echo sounder on range in the transducer nearfield, as measured in the freshwater tank. Superimposed on the data is the response as modeled from the transducer geometry.

carried out in advance of the acoustic measurements. Immediately following a review of the acoustic target-positioning operation, which also accomplishes mapping of the central beam pattern, measurement of the system response and stability, and choice of the receiver gain function, are described. Measurement of system directionality in each of several dimensions is then addressed. Measurement of the noise environment and system gain are also reviewed.

A. Preparations

In the case of multibeam sonars consisting of separate transmitting and receiving arrays, it may be necessary or advantageous to orient the arrays so that their acoustic axes intersect at the target range. This is particularly important for reducing effects of parallax in the case of nearfield calibration measurements to be applied in the transducer farfield. An easy method for adjusting the relative array orientations is to place shims between the bolts and mounting plate at the far end of one of the transducer arrays.

Another, essential operation in preparing for a calibration exercise is wetting the transducer faces and target to preclude formation of air bubbles on the respective surfaces.⁴⁷ This can be done by washing with a surfactant, for example, a simple household liquid detergent.

Recording the hydrographic state over the range of depths spanned by the transducer and target is essential in any calibration exercise. This is important for excluding stratification and other inhomogeneities in the water column as a cause of refractive effects during a calibration. At sufficiently near ranges, the effect of these on the propagation path between transducer and target may be entirely negligible, but they may be a limiting factor in others. In field applications over longer ranges, especially with horizontal or near-horizontal orientations, refractive effects, which are determined by the hydrography, may be substantial.

Beamforming is essential for the rapid and timely analysis of data. If desired or necessary, it may be performed digitally by software on the basis of the full set of individual

element echo signals and detailed knowledge of the receiving aperture geometry. Alternatively, the output signals from the sonar beamformer may be recorded and used directly in the data analyses.

B. Target positioning and measurement of central polar beam pattern

The basic procedure for relative positioning of the target in the transducer beam has already been outlined in Sec. IV A. It is reduced to a prescription in the following, assuming that the echo measurements are made in or close to the horizontal plane.

The standard target is suspended at a fixed distance from the centered transducer at the nominal transducer depth, specifically, the depth of the center of the transducer array or arrays. The target echo is registered. The target is now lowered by a fixed increment corresponding to some fraction of the nominal beamwidth of the transducer, say one-tenth. The echo measurement at this new position is repeated. The target is lowered again by the same fixed increment and its echo is recorded. This process is repeated until the first null is passed, which can be determined by observing a rise in the echo amplitude immediately following a steady decrease. The target is returned to the first measurement depth and the original measurement is repeated. The target is now raised by the fixed increment, and the echo is recorded. This process is repeated until the first null is again passed.

Beamforming is performed, if necessary, and the peak amplitude from the full set of beams is selected and plotted. The central equatorial plane of sensitivity is defined generally by the maximum in the plotted function. This may be determined by simple visual inspection if unambiguous, otherwise by fitting a quadratic or other nonlinear function to the data.

The exact position of the central axis can now be determined by rotating the transducer and observing the peak echo in the vicinity of the mechanically determined central axis. If the signal-to-noise ratio is sufficiently high, phase differences can be used to determine the angle of greatest sensitivity, which defines the central axis.⁴⁸ In either case, this refined determination of the central axis can reveal possible mechanical or electrical offsets in the central acoustic axis of the sonar. If the central axis deviates very much from the assumed axis, the measurements of polar directionality can be repeated to ensure their precision.

In the course of finding the central equatorial plane, the central polar beam pattern is measured. Noncentral polar beam patterns can be determined in a similar way.

C. System response, stability, and choosing receiver gain function

Having found the central plane of sensitivity, the system response can be determined for particular array orientations by observing the echo. By relating the measured output, e.g., amplitude, to the corresponding input signal and backscattering properties of the standard target, the transfer function of the system can be measured. This may involve analyzing the

frequency content of the transmitted signal and frequency dependence of the target backscattering cross section.

By repeating the measurements for the same transducer orientation and target position over a period of time, the stability of the system can be gauged. This can be quantified by a conventional time-series analysis, with observation of possible changes in the mean and variance over time.

By repeating the measurements with different receiver gain functions, the dependence of the system response on this function can be determined. By fitting a linear function to the echo intensity data, a suitable time-varied-gain function for the measurements can be determined. The system performance at the selected gain function can be related to other gain functions by means of the repeated measurements of system response.

D. Equatorial-plane sensitivity

Given knowledge of the location of the central plane of sensitivity, the equatorial-plane sensitivity function can be determined by rotating the transducer array or arrays over the angular span of interest. At each individual orientation, the several beam outputs are examined and the peak amplitude is selected. This process is repeated over the full angular span. The function defined by the peak output versus angle constitutes the equatorial-plane sensitivity.

E. Central-region directionality in two dimensions

The measurements of the equatorial-plane sensitivity function are now repeated at each of a succession of target depths. As in the target-positioning operation (Sec. V B), the target is moved systematically from the central plane past the first null below the plane, then from the central plane past the first null above the plane. Knowing the full target depth range spanning the main lobe, the measurements can also be performed in a single direction from one null to the other. By measuring the sensitivity function at each of these discrete target depths, the central-region directionality can be mapped in two dimensions.

F. Directionality of individual beams

The directionality of individual beams is measured in the central plane of sensitivity, as in Sec. V D, and over the central region of sensitivity, as in Sec. V E. The basic beamformed outputs are identical. Instead of selecting peak amplitudes from the range of beams for each transducer orientation, the amplitudes are associated with the respective individual beam. The individual beam patterns can thus be mapped. Individual beams are conveniently identified by their order in the series or by their nominal angle of peak sensitivity.

G. Nearfield response along acoustic axis

In cases where the multibeam sonar is to be used in the transducer nearfield or where the sonar is to be used in the farfield but can only be calibrated in the nearfield, the nearfield response should be measured. This is done by moving either the transducer or standard target along the acoustic

axis and measuring the echo response at discrete positions. The beamformed results are plotted against the transducer-target range.

The measurements are used for comparison against computations based on the transducer geometry and transmit frequency. Confirmation of the computational results provides support for the model, which can then be used to extrapolate either to greater or lesser ranges, depending on the intended sonar application.

H. Noise measurements

In all of the foregoing protocols, the echo from a standard target was generally measured. These measurements should be accompanied by similar measurements made in the absence of the target, hence of volume reverberation.

The reverberation measurements might be performed conveniently immediately before or immediately after a series of measurements with a standard target. While echoes from transient targets such as schools of juvenile fish passing through the sea well might not be recorded, the general level of background reverberation would be. Such measurements are important for describing the reverberation background and for quantifying the signal-to-noise ratio present during the target measurements. If the reverberation level is sufficiently high, another target with greater target strength might be chosen to ensure a better signal-to-noise ratio.

I. Electronic measurements

While purely electronic measurements are inadequate for determining the overall transmit-receive sensitivity, because of neglect of the transducer, they are nonetheless very useful for determining how signals are transferred in the receiver. The basis of the measurements is that of the time-amplitude-frequency measurement method.⁴⁹

The transducer is decoupled from the receiver. Signals of known amplitude and time delay are introduced into each channel, and the output is observed over a range of amplitudes sufficient to test the dynamic range and application of gain to each signal. The measurement sequence is repeated for each transducer element.

A complication with this simple notion is that of multiplexing of the individual transducer element signals before transmission to the beamforming processor, i.e., beamformer. Introduction of a signal of known amplitude and time delay is thus made to a multiplexer, with reception at the end of a cable. A special manufacturer-built unit can be entertained for measurement of the overall receiver sensitivity function with decoupled transducer. Such a capability would allow measurement of the electronic system to be undertaken independently of that of the combined acoustic and electronic system.

VI. DISCUSSION

By means of the calibration protocols described here, which have been developed through actual testing in nearly every instance, the performance of multibeam sonars can be defined. The general system sensitivity can be expressed through transfer characteristics and measures of directional-

ity. In addition, the stability can be quantified, and the connection between nearfield and farfield characteristics can be established.

The standard-target calibration method, which has been extended to multibeam sonar, is useful for its absolute character. Multibeam sonars calibrated in this way can be used in quantitative applications, for example, to estimate the numerical density of fish aggregations through echo-counting or echo-integration techniques. Visualization of fish aggregations by multibeam sonar, which is more often accomplished by uncalibrated systems, also benefits from a standard-target calibration, for variations in angular sensitivity can be compensated, giving a more accurate impression of the relative concentration densities.

Another practical advantage of calibration is making possible intercomparability of measurements with different multibeam sonars. Whether the several systems operate at the same or different frequencies, the measurements can be performed in an absolute sense. Given knowledge of the backscattering properties of the target, including the frequency dependence of target strength, if necessary the several acoustic measurements can be converted to biological or other physical measures of density, opening the possibility of direct comparison.

While the present work is directed to the calibration of multibeam sonars used for water-column measurements, it can also be applied to bathymetric multibeam sonars. In addition to measuring the range to the seafloor over a swath, the backscattering properties of the same can be measured and expressed in absolute units of surface scattering. Such measurements are essential for characterizing the seafloor and detecting changes in this, for example, to observe possible changes to the seafloor over time, as due to natural or man-made disturbances.

By quantifying the performance of multibeam sonar, conformity of manufactured units to their specifications can be confirmed. In addition, the performances of different sonars can be compared.

A number of outstanding problems in multibeam sonar calibration are identified. One is to generalize Eq. (1) to arbitrary operating conditions and ranges spanning the nearfield–farfield transition region of the sonar transducer. Another is to identify the causes of variability, for example, those encountered in measurements of system stability, already noted in Sec. IV C and observed in Fig. 8, as well as in Ref. 50. A third problem is to realize the electronic measurements described in Sec. VI in a rapid, automated manner. This could be done through a computer-controlled time-amplitude-frequency device. Another problem is extending the calibration protocols to multibeam sonars as configured for use, for example, as mounted within a sonar dome on the hull of a research vessel or on a fixed structure.

It is appreciated that rigorous calibration of multibeam sonars as configured for use may be difficult. In such cases, confirmation of the performance rather than measurement of the performance can be entertained. The problem of validation is thus identified. Derivative procedures, possibly with relaxed tolerances, could be very useful.

VII. SUMMARY

The goal of the project documented here was to develop convenient protocols for the calibration of multibeam sonar. The protocols that have been developed, based on the standard-target method, will enable absolute physical measurements to be made with the sonar. Addressing identified outstanding issues will enable the protocols to be extended efficiently to sonars as configured for use. It will also enable the overall accuracy of echo signal processing in the receiver to be quantified.

ACKNOWLEDGMENTS

The following are thanked for their participation in one or more of the calibration trials: B. Doherty, K. W. Doherty, S. P. Liberatore, and M. Parmenter, Woods Hole Oceanographic Institution; P. Lavoie, A. McLeod, and J. Millar, University of New Hampshire; Y. Li, formerly at the University of New Brunswick, Canada; G. D. Melvin, Department of Fish and Oceans, Canada; W. Michaels and D. A. Perry, Northeast Fisheries Science Center, Woods Hole. A. Bowen and D. R. Yoerger, Woods Hole Oceanographic Institution, are thanked for allowing their 6000-m-depth-rated SM2000/200-kHz multibeam imaging sonar to be used in an early trial. Several multibeam sonar manufacturers and their representatives are thanked for lending or arranging loans of sonars for use in developing the calibration protocols, as well as for participating in various trials. These include R. Asplin and C. Smith, Kongsberg Simrad Mesotech, Port Coquitlam, B.C., Canada; J. Condiotty, Simrad, Inc.; and S. Johnson and R. Morton, RESON, Inc. L. Bjørnø, J. Steensrup, and P. K. Eriksen are thanked for providing data on the RESON Sea-Bat 8101 sonar array. J. Doucette is thanked for assistance in preparing figures; S. Barkley, A. Norton, and M. Parmenter are thanked for assistance in assembling the manuscript. Support by the National Science Foundation through Award No. OCE-0002664, NOAA through Grant No. NA97OG0241, and the Cooperative Institute for Climate and Ocean Research (CICOR) through NOAA Contract No. NA17RJ1223 is acknowledged. This is Woods Hole Oceanographic Institution Contribution No. 11178.

¹H. Medwin and C. S. Clay, *Fundamentals of Acoustical Oceanography* (Academic, Boston, 1998).

²R. B. Mitson, *Fisheries Sonar* (Fishing News Books, Farnham, Surrey, England, 1983).

³O. A. Misund and A. Aglen, "Swimming behavior of fish schools in the North Sea during acoustic surveying and pelagic trawl sampling," *ICES J. Mar. Sci.* **49**, 325–334 (1992).

⁴F. Gerlotto, P. Fréon, M. Soria, P. H. Cottais, and L. Ronzier, "Exhaustive observations of the 3D schools structure using multibeam sidescan sonar: potential use for school classification, biomass estimation, and behavior studies," Council Meeting of the Int. Council Explor. Sea 1994/B:26 (1994).

⁵F. Gerlotto, M. Soria, and P. Fréon, "From two dimensions to three: the use of multibeam sonar for a new approach in fisheries acoustics," *Can. J. Fish. Aquat. Sci.* **56**, 6–12 (1999).

⁶M. Soria, P. Fréon, and F. Gerlotto, "Analysis of vessel influence on spatial behaviour of fish schools using a multi-beam sonar and consequences for biomass estimates by echo-sounder," *ICES J. Mar. Sci.* **53**, 453–458 (1996).

⁷O. A. Misund, "Dynamics of moving masses: variability in packing den-

- city, shape, and size among herring, sprat and saithe schools,” ICES J. Mar. Sci. **50**, 145–160 (1993).
- ⁸M. T. Hafsteinnsson and O. A. Misund, “Recording the migration behavior of fish schools by multibeam sonar during conventional acoustic surveys,” ICES J. Mar. Sci. **52**, 915–924 (1995).
- ⁹D. G. Reid (ed.), “Report on echo trace classification,” ICES Coop. Res. Rep. (238) (2000).
- ¹⁰L. Mayer, Y. Li, and G. Melvin, “3D visualization for pelagic fisheries research and assessment,” ICES J. Mar. Sci. **59**, 216–225 (2002).
- ¹¹T. C. Gallaudet and C. P. de Moustier, “High-frequency volume and boundary acoustic backscatter fluctuations in shallow water,” J. Acoust. Soc. Am. **114**, 707–725 (2003).
- ¹²L. Nøttestad and B. E. Axelsen, “Herring schooling manoeuvres in response to killer whale attacks,” Can. J. Zool. **77**, 1540–1546 (1999).
- ¹³B. E. Axelsen, T. Anker-Nilssen, P. Fossum, C. Kvamme, and L. Nøttestad, “Pretty patterns but a simple strategy: predator-prey interactions between juvenile herring and Atlantic puffins observed with multibeam sonar,” Can. J. Zool. **79**, 1586–1596 (2001).
- ¹⁴K. Benoit-Bird and W. Au, “Hawaiian spinner dolphins aggregate midwater food resources through cooperative foraging,” J. Acoust. Soc. Am. **114**, 2300 (2003).
- ¹⁵T. Weber, D. Bradley, R. L. Culver, and A. Lyons, “Inferring the vertical turbulent diffusion coefficient from backscatter measurements with a multibeam sonar,” J. Acoust. Soc. Am. **114**, 2300 (2003).
- ¹⁶C. D. Jones, “Water-column measurements of hydrothermal vent flow and particulate concentration using multibeam sonar,” J. Acoust. Soc. Am. **114**, 2300–2301 (2003).
- ¹⁷R. J. Urick, *Principles of Underwater Sound*, 3rd ed. (McGraw–Hill, New York, 1983).
- ¹⁸K. G. Foote, “Optimizing copper spheres for precision calibration of hydroacoustic equipment,” J. Acoust. Soc. Am. **71**, 742–747 (1982).
- ¹⁹K. G. Foote, “Maintaining precision calibrations with optimal copper spheres,” J. Acoust. Soc. Am. **73**, 1054–1063 (1983).
- ²⁰K. G. Foote, H. P. Knudsen, G. Vestnes, D. N. MacLennan, and E. J. Simmonds, “Calibration of acoustic instruments for fish density estimation: a practical guide,” ICES Coop. Res. Rep. 144 (1987).
- ²¹D. Chu, K. G. Foote, and L. C. Hufnagle, Jr., “Measurement of multibeam sonar directivity patterns,” Proc. MTS/IEEE Oceans 2002 Conf., pp. 1411–1414 (2002).
- ²²N. A. Cochrane, Y. Li, and G. D. Melvin, “Quantification of a multibeam sonar for fisheries assessment applications,” J. Acoust. Soc. Am. **114**, 745–758 (2003).
- ²³D. Chu, K. G. Foote, L. C. Hufnagle, Jr., K. C. Baldwin, L. A. Mayer, and A. McLeod, “Calibrating a 90-kHz multibeam sonar,” Proc. MTS/IEEE Oceans 2003 Conf., pp. 1633–1636 (2003).
- ²⁴K. G. Foote, D. Chu, K. C. Baldwin, L. A. Mayer, A. McLeod, L. C. Hufnagle, Jr., J. M. Jech, and W. Michaels, “Protocols for calibrating multibeam sonar,” J. Acoust. Soc. Am. **114**, 2307 (2003).
- ²⁵G. C. Gaunard, “Sonar cross section of bodies partially insonified by finite sound beams,” IEEE J. Ocean. Eng. **OE-10**, 213–230 (1985).
- ²⁶A. D. Poularikas, “Fourier transforms and the fast Fourier transform,” in *The Electrical Engineering Handbook*, 2nd edition, edited by R. C. Dorf (CRC, Boca Raton, FL, 1997), pp. 287–298.
- ²⁷K. G. Foote and D. N. MacLennan, “Comparison of copper and tungsten carbide calibration spheres,” J. Acoust. Soc. Am. **75**, 612–616 (1984).
- ²⁸K. G. Foote, P. R. Atkins, C. C. Bongiovanni, D. T. I. Francis, P. K. Eriksen, M. Larsen, and T. Mortensen, “Measuring the frequency response function of a seven-octave-bandwidth echo sounder,” Proc. Inst. Acoust. **21(1)**, 88–95 (1999).
- ²⁹K. G. Foote, D. T. I. Francis, H. Furset, and H. Hobæk, “Spheres for calibrating high-frequency broadband echo sounders,” Acust. Acta Acust. **85**, S186–S187 and J. Acoust. Soc. Am. **105**, 1112–1113(A) (1999).
- ³⁰D. L. Folds, “Target strength of focused liquid-filled spherical reflectors,” J. Acoust. Soc. Am. **49**, 1596–1599 (1971).
- ³¹R. H. Wallace, H. V. Hillery, G. R. Barnard, B. M. Marks, and C. M. McKinney, “Experimental investigation of several passive sonar targets,” J. Acoust. Soc. Am. **57**, 862–869 (1975).
- ³²B. M. Marks and E. E. Mikeska, “Reflections from focused liquid-filled spherical reflectors,” J. Acoust. Soc. Am. **59**, 813–817 (1976).
- ³³D. L. Folds and C. D. Loggins, “Target strength of liquid-filled spheres,” J. Acoust. Soc. Am. **73**, 1147–1151 (1983).
- ³⁴T. R. Hammar, K. G. Foote, and S. P. Liberatore, “Advances in developing a high-frequency sonar calibration facility,” Proc. MTS/IEEE Oceans 2003 Conf., pp. 1622–1624 (2003).
- ³⁵D. M. Deveau, “Fabrication, testing, and installation of fluid-filled target spheres for AUTECS’s passive sonar calibration system,” Proc. MTS/IEEE Oceans 2000 Conf., pp. 1481–1487 (2000).
- ³⁶J. J. Faran, Jr., “Sound scattering by solid cylinders and spheres,” J. Acoust. Soc. Am. **23**, 405–418 (1951).
- ³⁷R. Hickling, “Analysis of echoes from a solid elastic sphere in water,” J. Acoust. Soc. Am. **34**, 1582–1592 (1962).
- ³⁸R. R. Goodman and R. Stern, “Reflection and transmission of sound by elastic spherical shells,” J. Acoust. Soc. Am. **34**, 338–344 (1962).
- ³⁹K. W. Doherty, T. R. Hammar, and K. G. Foote, “Transducer mounting and rotating system for calibrating sonars in a sea well,” Proc. MTS/IEEE Oceans 2002 Conf., pp. 1407–1410 (2002).
- ⁴⁰K. C. Baldwin, L. Mayer, A. McLeod, and J. Millar, “Acoustic transducer calibration system,” Proc. MTS/IEEE Oceans 2003 Conf., pp. 2093–2099 (2003).
- ⁴¹R. F. Duncan and F. S. Blanchard, *A Cruising Guide to the New England Coast*, 5th ed. (Dodd, Mead, New York, 1962).
- ⁴²C. W. Horton, Sr., *Signal Processing of Underwater Acoustic Waves* (U.S. Government Printing Office, Washington, DC, 1969).
- ⁴³W. B. Davenport, Jr., “Signal-to-noise ratios in band-pass limiters,” J. Appl. Phys. **24**, 720–727 (1953).
- ⁴⁴H. Bodholt, H. Nes, and H. Solli, “A new echo-sounder system,” Proc. Inst. Acoust. **11(3)**, 123–130 (1989).
- ⁴⁵N. R. Draper and H. Smith, *Applied Regression Analysis* (Wiley, New York, 1966).
- ⁴⁶K. G. Foote, “Acoustic sampling volume,” J. Acoust. Soc. Am. **90**, 959–964 (1991).
- ⁴⁷L. E. Ivey and C. M. Thompson, “Underwater transducer wetting agents,” J. Acoust. Soc. Am. **78**, 389–394 (1985).
- ⁴⁸D. Chu, K. C. Baldwin, K. G. Foote, Y. Li, L. A. Mayer, and G. D. Melvin, “Multibeam sonar calibration: target localization in azimuth,” Proc. MTS/IEEE Oceans 2001 Conf., pp. 2506–2510 (2001).
- ⁴⁹H. P. Knudsen, “T-A-F: Time-amplitude-frequency: A special electronic unit for measuring the TVG function in research echo-sounders,” Fiske-ridir. Skr., Ser. Havunders. **17**, 529–541 (1985).
- ⁵⁰D. Chu, K. G. Foote, T. R. Hammar, L. C. Hufnagle, Jr., and J. M. Jech, “Calibrating a 90-kHz multibeam sonar: illustrating protocols,” Proc. MTS/IEEE Oceans 2004 Conf., pp. 438–442 (2004).

Multi-mode Lamb wave tomography with arrival time sorting

Kevin R. Leonard and Mark K. Hinders

Department of Applied Science, College of William and Mary, Williamsburg, Virginia 23187-8795

(Received 22 July 2004; revised 13 January 2005; accepted 14 January 2005)

Lamb wave tomography has been shown to be an effective nondestructive evaluation technique for platelike structures. A series of pitch-catch measurements between ultrasonic transducers can be taken from different orientations across the sample to create a map of a particular feature of interest such as plate thickness. Most previous work has relied solely on the first arriving mode for the time-of-flight measurements and tomographic reconstructions. The work described here demonstrates the capability of the Lamb wave tomography system to generate accurate reconstructions from multiple modes. Because each mode has different through-thickness displacement values, each is sensitive to different types of flaws, and the information gained from a multi-mode analysis can improve understanding of the structural integrity of the inspected material. However, one of the problems with the extraction of multi-mode arrival times is that destructive interference between two modes may cause one of the modes to seemingly disappear in the signal. The goal of the sorting algorithm presented in this work is to try and counteract this problem by using multiple frequency scans—also known as frequency walking—to sort the arrival times into their correct mode series. © 2005 Acoustical Society of America.

[DOI: 10.1121/1.1867792]

PACS numbers: 43.35.Pt [TDM]

Pages: 2028–2038

I. INTRODUCTION

Lamb waves are useful for the inspection of a variety of platelike structures. If Lamb wave measurements are made for a large number of specific pitch-catch transducer positions surrounding a region of interest, then an image of that region can be reconstructed tomographically. The reconstruction is an easily interpretable quantitative map of a parameter of interest, e.g., thickness loss due to corrosion.

Key to any practical Lamb wave tomography (LWT) scheme is a suite of robust and reliable algorithms that automatically interprets the digitized Lamb waveforms. In bulk wave ultrasonics gating and peak-detection schemes are usually adequate. With guided waves, however, sophisticated signal processing is required in order to identify the various Lamb wave modes in the digitized waveforms and then to extract the feature(s) of interest from each mode as input for the reconstruction algorithm. Fairly small errors in this mode extraction step usually spoil the reconstruction completely. Also, for a practical nondestructive evaluation (NDE) system these algorithms must run in real time on portable hardware, and must be able to deal with millions of digitized waveforms per minute.

Early Lamb wave tomography work by Hutchins *et al.*,^{1–5} Achenbach,⁶ and Degertekin⁷ used a parallel projection geometry with the velocity and/or attenuation of Lamb waves as input for the tomographic reconstructions. McKeon and Hinders^{8–12} implemented this also, and then investigated a “crosshole” tomographic scheme similar to a preliminary pipe inspection study done by Hilderbrand *et al.*¹³ that has many practical advantages for aircraft NDE. Malyarenko and Hinders¹⁴ subsequently compared fan beam and double crosshole geometries, and found the latter to be superior. A secondary conclusion of that work is that the iterative reconstruction algorithms from the seismological literature seem

to be better suited to Lamb wave NDE applications than are the convolution-backprojection type of algorithms from the medical imaging literature. Malyarenko *et al.*^{15–17} have implemented a curved ray tracing technique that accounts for scattering or refractive media and improves the technique’s accuracy in sizing flaws.

Reconstructions that only use the time-of-flight information for the first arriving mode at a predetermined frequency may be discarding a wealth of information in the rest of the signal. The various Lamb wave modes interact with flaws differently because of their different through-thickness displacement properties, and certain modes only interact weakly with certain flaws. Therefore, delineating the different mode arrivals could significantly enhance the capabilities of the LWT system.

II. EXPERIMENTAL SETUP

In our current laboratory system we use a pair of independently scanning pitch-catch transducers that are attached to linear slider screws and are moved back and forth by computer controlled stepper motors. The data acquisition equipment and scanning system consist of a Matec TB-1000 toneburst pulser-receiver board and a Gage CS8012A 12-bit A/D board contained within a PC running Linux. The stepper motors are under computer control via a Velmex 9000 controller. Figure 1 shows a schematic of the scanning system. For the results presented in this work, a pair of 0.5-in.-diam, 1-MHz longitudinal contact transducers are used for both transmission and reception of the Lamb wave signals. The total scanned area is 20×20 cm², and the step size between scanning locations is 2 mm. As the sample is scanned, the waveforms for each projection are digitized, stacked into a data file and stored for subsequent data analysis. For each tomographic projection, the transmitting transducer steps

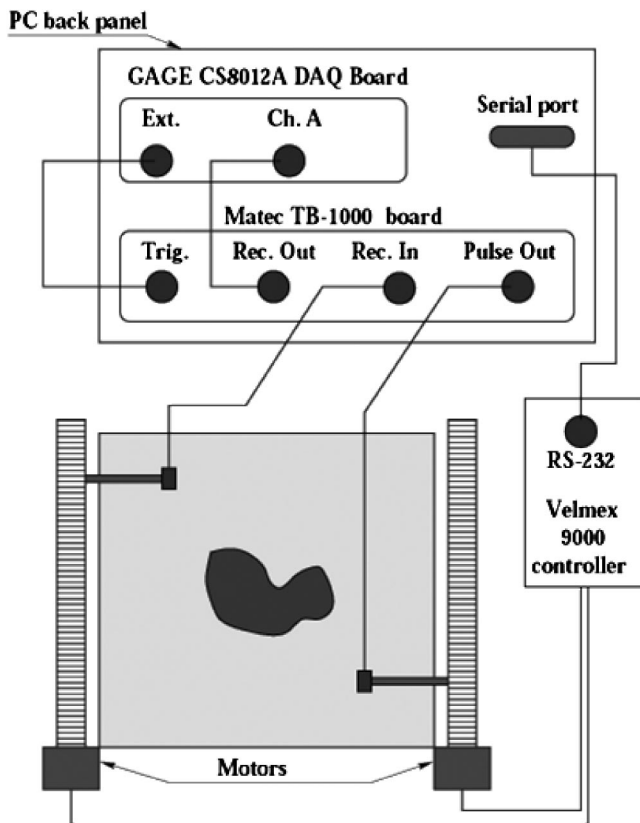


FIG. 1. Schematic of the Lamb wave tomography (LWT) scanner.

along one edge of the perimeter of a rectangular region of interest while the receiving transducer steps through all the positions on a different edge. This gives us the large number of crisscross ray paths necessary for tomographic reconstruction, and mimics a situation where a four-legged “picture-frame” array of multiplexed transducers is used to surround the area to be inspected.

In order to evaluate the effectiveness of the arrival time extraction algorithm, an intermediate representation of the tomographic data is needed. We have found that arrival time scatter plots work well for this purpose. For example, Fig. 5(a) is an arrival times scatter plot for a complete projection of a flawless aluminum plate. The x axis is called the pitch-catch, or waveform, number. In the LWT scanner, the transmit transducer remains stationary while the receive transducer steps through the 100 scan positions (the total scan area is 20×20 cm² and the step size for both the transmit and receive transducers is 2 mm). Once the receive transducer has stepped through all the positions the transmit transducer steps a single position. The receive transducer then steps through its 100 positions in reverse until it reaches its original starting position. Again, the transmit position steps once and the receive transducer then steps through all of its positions. This process is repeated until the transmit transducer has stepped through all 100 positions. Therefore, the pitch-catch number—or waveform number—is $(transmit_position - 1) * 100 + receive_position$. This also explains why the arrival times as a whole seem to decrease and then increase again in a symmetrical pattern. This occurs as the transmit transducer moves through all of its

scanning positions. For a clean plate, this pattern should be symmetrical and regular as shown. Flaws appear in these representations as deviations from the symmetrical flawless patterns. This intermediate representation allows us to make sure that the data were recorded properly and it also can serve as a measure of the performance of different time extraction algorithms since errors in the arrival times show up as extreme deviations from the expected pattern.

III. WAVELET ANALYSIS AND ARRIVAL TIME EXTRACTION

The method used previously to extract the arrival times of the fastest Lamb wave modes is based on enveloping the signal and is described in previous work.^{10–12,14–20} It is a straightforward and computationally inexpensive method, which has been shown to perform better than other more complicated and computationally expensive time-frequency methods.¹⁶ However, even though the simple thresholding arrival time extraction algorithm is effective, certain limitations still exist. For instance, if the amplitude of the mode of interest drops below the threshold level—as it often does after it interacts with a flaw—then the extraction algorithm will pick up a later mode. Furthermore, it is incapable of providing arrival time information for multiple modes.

Malyarenko¹⁶ explored various time-frequency methods to try and improve the arrival time extraction methods. However, Wigner-Ville distributions, short-time Fourier transforms (STFTs), and positive distributions did not sufficiently improve the accuracy of the arrival times. In work recently done by Hou *et al.*,^{21,22} wavelet analysis was explored as a better solution to the current arrival time extraction problem. Wavelets have become popular in many different types of signal and image processing techniques. They are particularly advantageous for both the denoising and compression of signals.²³

The continuous wavelet transform is defined as:

$$W_g(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) g_{a,b}(t) dt, \quad (1)$$

where $g_{a,b}(t) = g^* \left(\frac{t-b}{a} \right)$.

$W_g(a,b)$ is the wavelet coefficient for the translation b and the scale a . $g_{a,b}$ is the scaling function which is calculated from the mother wavelet $g(t)$. By varying the translation and scale parameters, one can inspect the signal from different time and scale representations. This flexibility is a great advantage over the traditional Fourier analysis and STFT.

In our case, the signals being processed are discrete signals. The discrete wavelet transform (DWT) implements the CWT for discrete signals. Conceptually, the DWT can be viewed as a series of filters where the input signal is broken down into a low frequency approximation and a high frequency detail signal. The approximation signal can then be broken down further through the same process. The result will be an approximation signal with a number of detail signals. Each of the detail signals is a set of wavelet transform coefficients at the next finer scale. Denoising and compres-

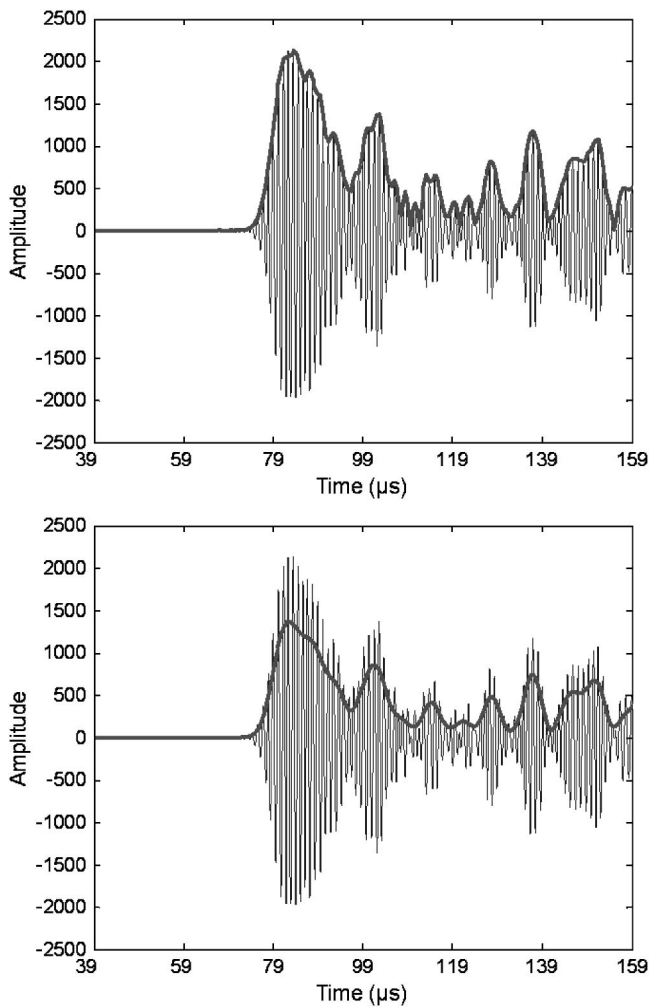


FIG. 2. Comparison of enveloping techniques: (a) wavelength peak enveloping and (b) wavelet enveloping.

sion are accomplished by setting a threshold value for the resulting coefficients. For most signals, the majority of information is contained within only a few coefficients. Furthermore, because the details are the high frequency components, by removing these coefficients, one can also reduce the noise in the signal—which is often higher in frequency and smaller in scale. This allows one to eliminate the noise, while keeping the information that is desired.

IV. SORTING ALGORITHM

The method presented here uses a version of the DWT to generate smooth envelopes for the Lamb wave signals. The peaks of these different envelopes are then used as the arrival times of the multiple Lamb wave modes. One of the problems with the extraction of multiple arrivals is that destructive interference between two modes may cause one of the modes to seemingly disappear in the recorded signal (see Fig. 4). The goal of the sorting algorithm presented in this section is to try and counteract this problem by using the frequency walk information described below to sort the arrival times into their correct mode series.

As discussed above, the current time extraction algorithm relies on the accurate enveloping of the signal. In the past this had been done using the peak values of each wave-

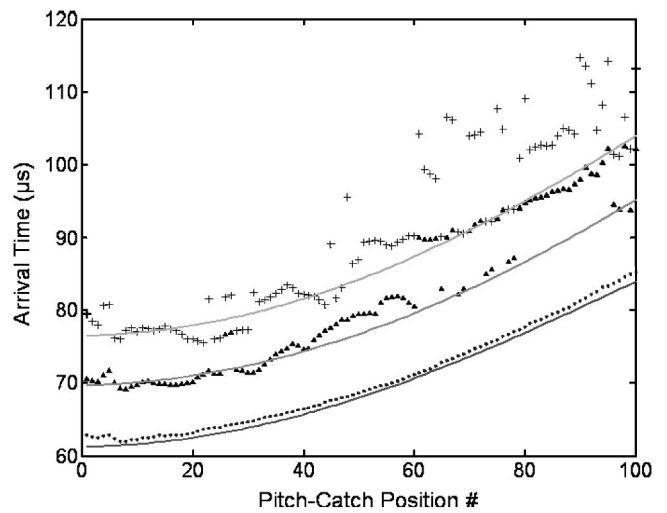


FIG. 3. Arrival times for the first three peaks of a clean aluminum plate at the first 100 receiver positions (points—peak 1; triangles—peak 2; crosses—peak 3). The solid lines are the fitted mode predictions. Notice that starting at position 61, the second peak arrives closer to the expected arrival of mode 3.

length. However, when trying to track more than just the first arriving mode, the unsmooth nature of this type of envelope made it difficult to detect subsequent modes. Therefore, wavelets were used to simultaneously denoise and envelope

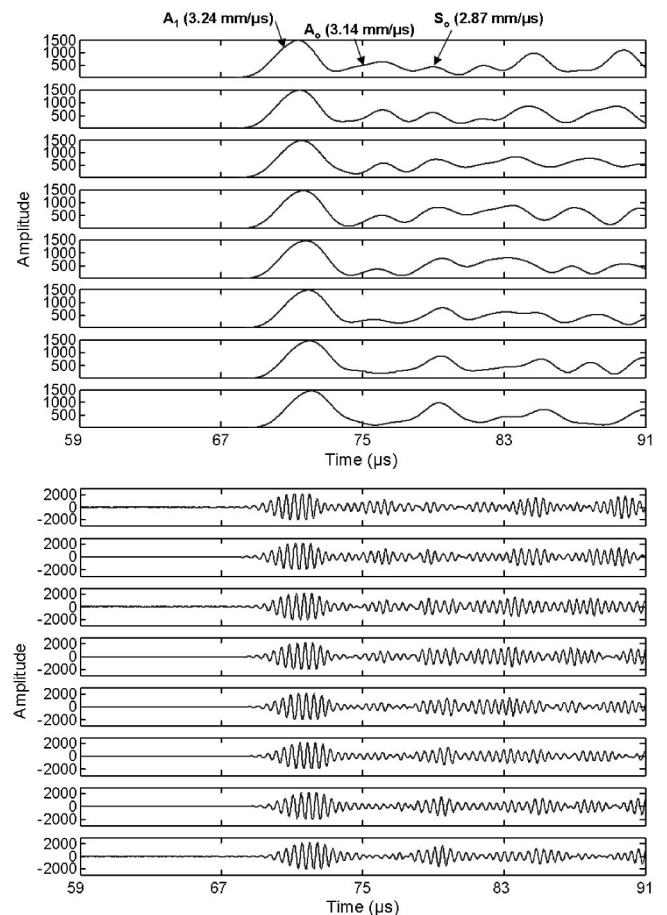


FIG. 4. The envelopes and waveforms of the recorded Lamb wave signals in the clean aluminum plate sample for receiver positions 55–62 (top to bottom). As the second and third wave packets interfere, the second peak disappears while the third peak increases in amplitude.

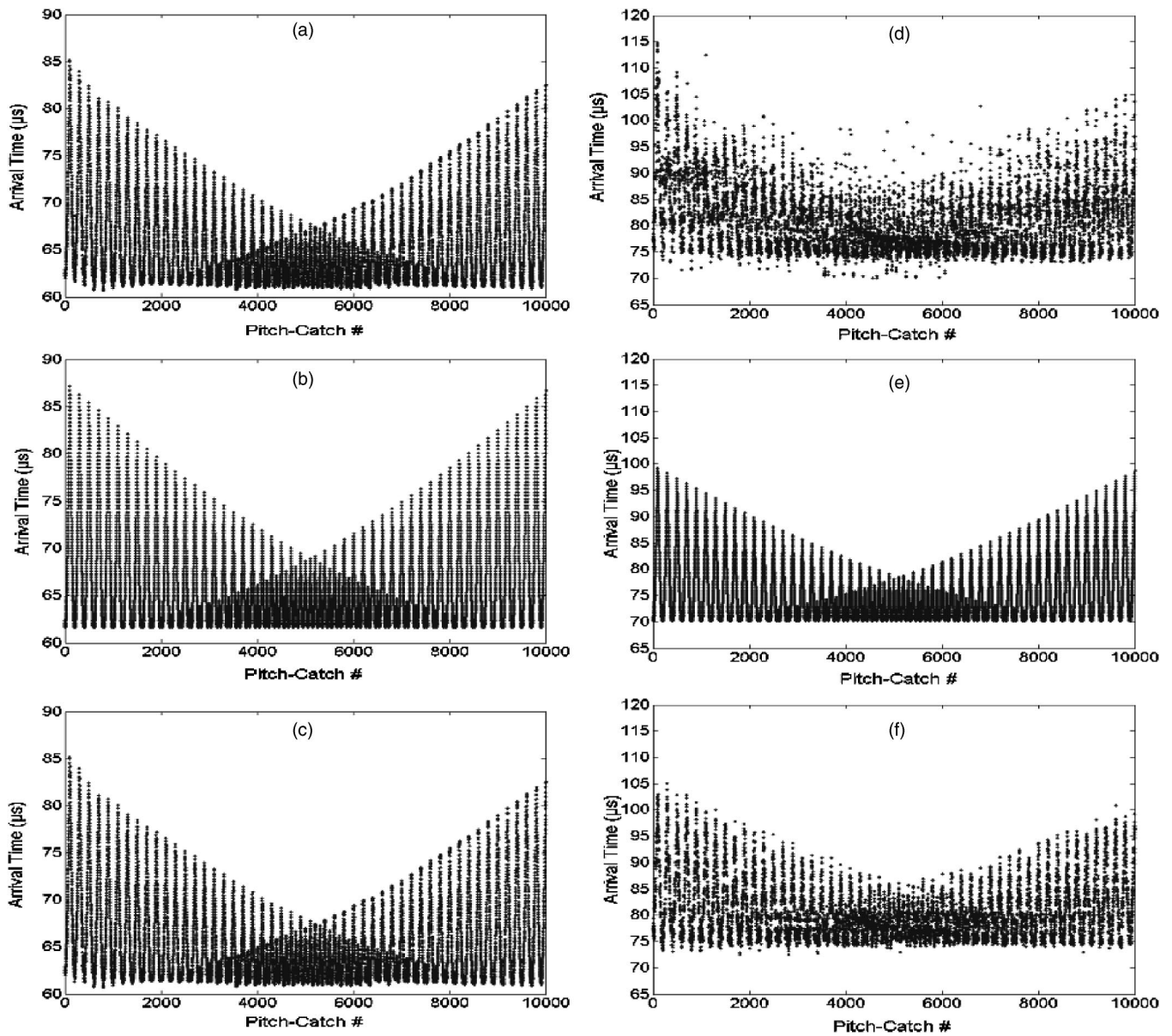


FIG. 5. Arrival time scatter plots for the A_1 and S_0 modes for a single projection of a clean aluminum plate. (a) Raw data of the first envelope peak arrival. (b) Theoretical data for first arriving mode: Group velocity of $A_1 = 3.25 \text{ mm}/\mu\text{s}$. (c) Sorted data for the first arriving mode. (d) Raw data of the third envelope peak arrival. (e) Theoretical data for third mode: Group velocity of $S_0 = 2.85 \text{ mm}/\mu\text{s}$. (f) Sorted arrival times for the third mode.

the signal with a pruning procedure²⁴ based on the discrete stationary wavelet transform.²⁵ Naturally, other methods could have been used to smooth the envelopes, but the following procedure integrates easily with the work recently performed by Hou *et al.*^{21,22} Ultimately, it is desired to combine the sorting algorithm described in this work with the wavelet detection algorithm developed by Hou *et al.* to form a robust multi-mode LWT system.

MatlabTM was used to compute the wavelet transform of the Lamb wave signals using an eight-level *coiflet3* decomposition. The *coiflet3* wavelet was chosen because it seemed to most closely resemble the wave packets we were trying to find the arrival times for. The absolute value of each signal was decomposed using the stationary wavelet transform. The first five levels of coefficients were zeroed and the remaining levels were inverse transformed to provide the envelope values. The ability of the wavelet transform to separate out the

high frequency noise and oscillations of the signal into the first five levels of detail is what makes this technique possible. Figure 2 shows a typical signal and compares the peak enveloping and wavelet enveloping techniques. It can be seen that the wavelet enveloping technique provides a smoother envelope and identifies fewer “false peaks.”

The wavelet enveloping allows us to extract the arrival times of the envelope peaks after the first arriving mode. Figure 3 shows the arrival times for the first three peaks in the recorded signal for the first transmitter position in a scan of a clean aluminum plate at 920 kHz. It can be seen that the first peak’s arrival time is extracted cleanly, but that the subsequent peak arrivals are much more scattered on the plot. The uncertainty in the arrival times of the later arriving peaks arises from constructive and destructive interference between dispersive modes. Figure 4 shows the signal envelopes and waveforms from receiver positions 55 to 62. At a

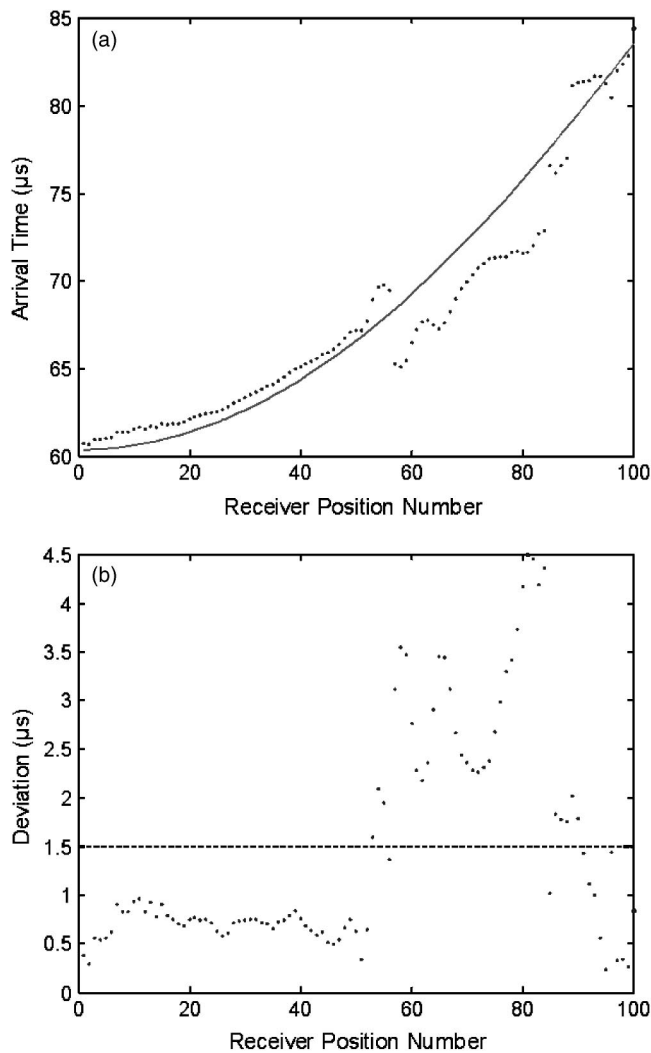


FIG. 6. (a) The first peak's arrival times for the first 100 points of projection 1 for the flat-bottomed hole sample (solid line is the polynomial fit of the arrival times for the first mode). (b) Deviations of the arrival times from the polynomial fit. Notice that a deviation threshold of $1.5 \mu\text{s}$ can be used to differentiate between the flawed and unflawed regions of the plate.

frequency thickness of $2.92 \text{ MHz}\cdot\text{mm}$, it is expected that the first arriving mode will be the A_1 mode with a group velocity of $3.24 \text{ mm}/\mu\text{s}$. It will be followed by the A_0 mode ($3.14 \text{ mm}/\mu\text{s}$) and then the S_0 mode ($2.87 \text{ mm}/\mu\text{s}$). The sequence of plots in Fig. 4 highlights how the envelope peaks combine because the individual dispersive modes are spreading as they propagate. The disappearance of the peak is less of an issue than the intermixing of the mode series. In order to generate a tomographic image from the later arriving mode arrival times, they need to be separated from each other. However, this problem becomes nontrivial because the data series jump between modes as certain modes interfere with each other. It is necessary to note that for accurate tomographic reconstructions, the measurements are designed to be taken in highly dispersive regions because the change in velocity due to thickness loss is the parameter of interest. Thus, in order to use the multi-mode information, a more robust time extraction algorithm is needed to overcome these interference problems.

For the clean plate, the arrival times can be sorted into

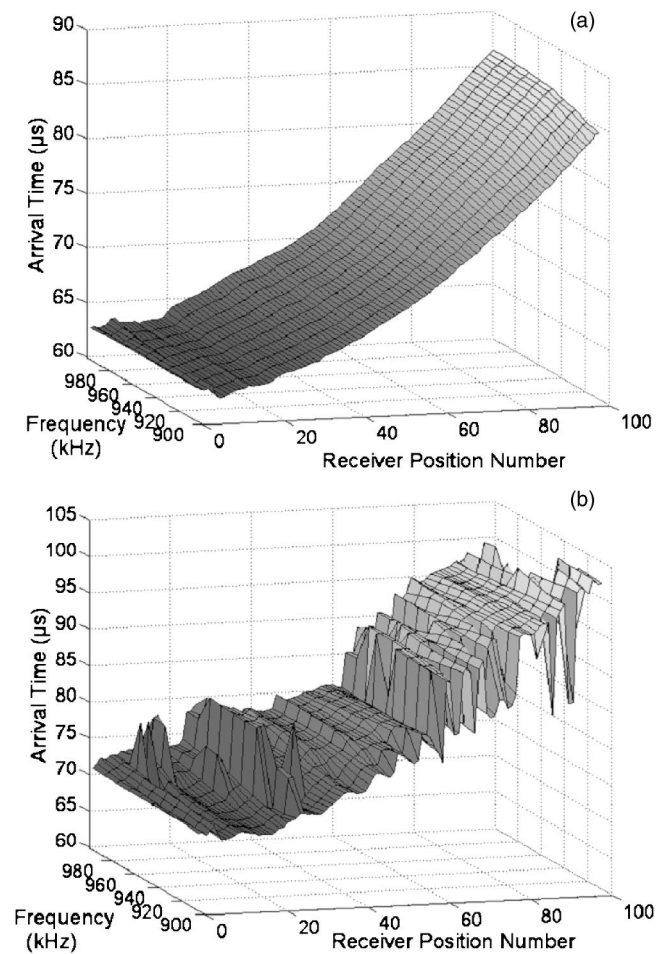


FIG. 7. Frequency walk surface plots for the arrival times of the first transmit position in a scan of a clean aluminum plate. Times recorded for ten evenly spaced frequencies between 900 and 990 kHz. (a) Arrival times for the first peak. (b) Arrival times for the second peak.

their correct series by fitting the individual series with a third-order polynomial fit. This provides a guess for where the missing points should lie and provides a measure to determine whether a mode has been missed. Figure 3 shows the arrival times for the first three envelope peaks and the polynomial fit for each data series. It can be seen that the polynomial fit effectively estimates where the mode should be and also shows how at receiver position 61, the second recorded peak arrival is actually the arrival time for the third mode.

In order to use the polynomial fit to sort the arrival times into their appropriate mode numbers, we compare the times to the fitted mode curves. For the clean plate, the arrival times are sorted into the series that corresponds to the fitted curve that they are closest to. If a particular mode is not present, we store the fitted value at that waveform number into the arrival time data as a placeholder for the tomographic reconstruction algorithms since we do not have enough information to determine when the actual arrival time is. Once the data has been sorted, the process can be iterated and we can refit the mode data and sort again to further improve the arrival times for the clean sample.

The sorting algorithm compared to both the theoretical and predicted arrival times reduced the mean square error

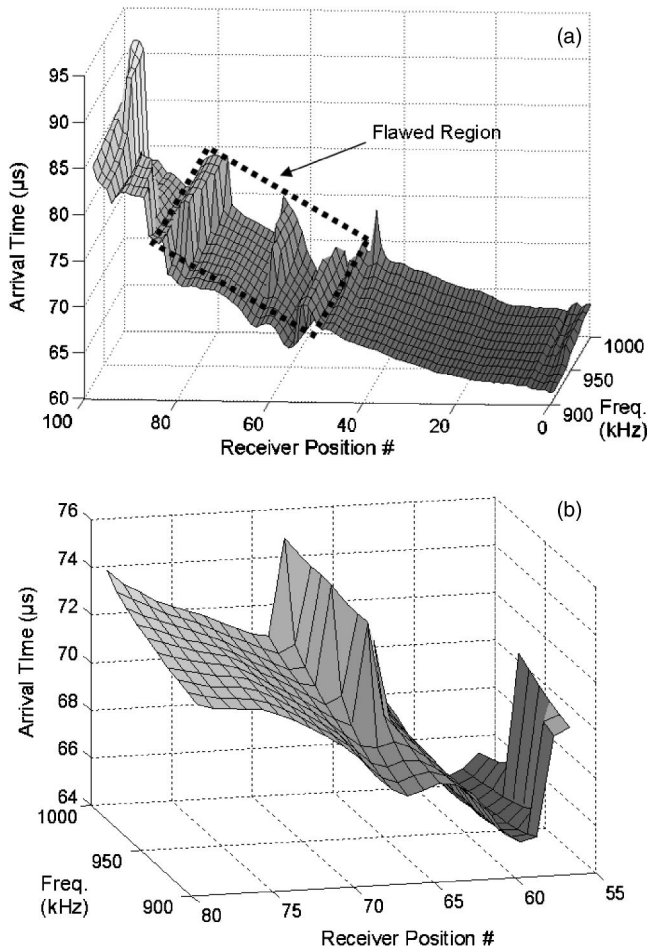


FIG. 8. (a) Frequency walk data for the first 100 arrivals of the first arriving mode in the flat-bottomed hole sample (see Fig. 5 for comparison with a clean plate). (b) Arrival times for the “flawed” region between receiver positions 55 and 80.

(MSE) of the measurement. The MSE for the unsorted arrival times compared to theoretical arrival times was 3.01, 6.32, and 8.26 μs for the first three modes, respectively. After sorting, the MSE was reduced to 3.01, 5.77, and 7.47 μs . Part of the large discrepancy between the theoretical and experimental arrival times of the modes is due to how the times were extracted. The arrival times are determined by the location of the envelope’s peak. Dispersion of the mode’s shape due to the bandwidth of the input signal can cause variations in the peak arrival. A more effective measure of arrival times has been found to be the location of the front of the envelope.¹⁶ However, for later arriving modes, this position is hard to obtain. In contrast, the MSE between the unsorted arrival times and the predicted arrivals (the fitted data) was 0.24, 2.24, and 3.48 μs for the first three modes and for the sorted times, the MSE was 0.24, 1.47, and 1.71 μs . Figure 5 shows the arrival time scatter plots before and after the sorting algorithm was applied for the A_1 and S_0 modes. As expected, the first arrival time is fairly accurate and is not affected by the sorting algorithm, but the third mode is significantly improved.

The previous algorithm works well for a clean sample, but will blur out flaws in the tomographic reconstructions because it assumes that any deviation from the predicted

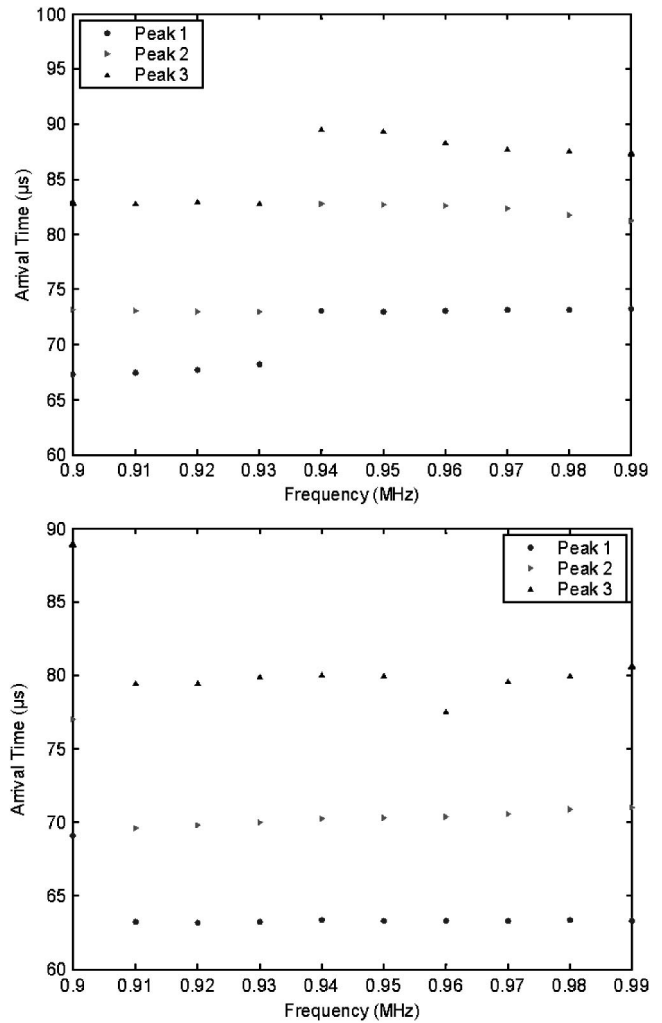


FIG. 9. (a) Arrival time versus frequency for the first three envelope peaks in signal 65. (b) Arrival time versus frequency for signal 598 where only the first frequency missed the first arriving mode.

curves is due to interference. To extend the usefulness of the sorting algorithm to flawed samples, more information is needed to determine whether the individual arrival times are due to a flaw, a missed mode, or a bad data point resulting from poor coupling or digitization. Because the first arrival is accurately detected and is not affected by the clean plate sorting algorithm, we can use the first arrival as a measure to see whether the subsequent arrivals can be seen as “clean” arrivals or need to be handled as “potential flaw” arrivals.

Figure 6(a) shows the first 100 points of a projection for a 3.17-mm-thick aluminum plate with a 2.2-in.-diam off-center flat-bottomed hole ($\sim 60\%$ thickness loss). From the geometry of this sample and the scanner, the Lamb wave data should begin to interact with the flaw around receiver position 55 for the first transmit position. The arrival time plot shows how the fastest mode arrives earlier in the flawed region because the frequency-thickness product has changed. It also should be noted that due to change in the frequency-thickness product, the second arrival also speeds up and arrives where the first mode would be located on a clean sample. This causes the sorting algorithm to choose the second arrival time as the first mode and thus discounts the actual “flaw” arrivals. In Fig. 6(b) the deviations of the ar-

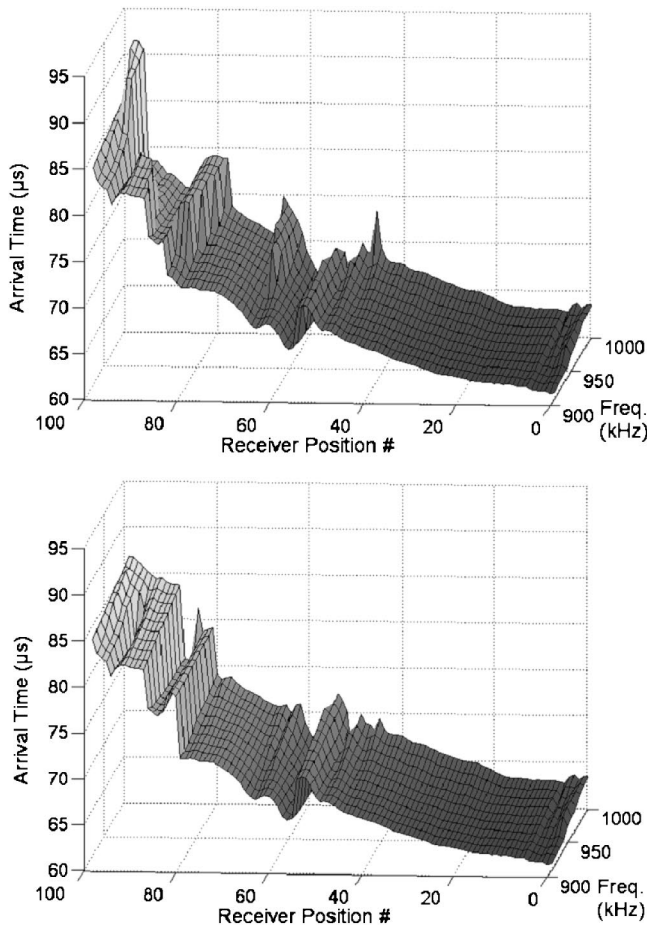


FIG. 10. Comparison of frequency walk data for the first arriving mode before sorting (top) and after sorting (bottom).

arrival times from the polynomial fit in the flawed sample are also shown. It can be seen that at position 55—where the Lamb wave data begin to interact with the flaw—the deviation in the measurement from the fitted data is greater than $1.5 \mu\text{s}$. Therefore, in order to delineate the suspected “clean” arrivals from suspected “flaw” arrivals, a deviation of $1.5 \mu\text{s}$ from the expected arrival was used as a threshold level.

Once the arrival times are able to be separated, it is still necessary to determine which mode they belong to and whether they are actual arrivals or erroneous data points. To gain more information, at each pitch-catch position the frequency was incremented at equally spaced intervals. The frequency was changed from 900 to 900 kHz at 10 kHz steps, and, for each frequency, the waveform was digitized and stacked into its corresponding data file. The frequency walking technique provides an extra dimension in order to solve the arrival time sorting problem. As the frequency is slightly varied, the arrival times of the different lamb wave modes should also change. By selecting a particular region on the group velocity dispersion curves, this can provide more information about which modes are actually being recorded. The 900- to 990-kHz region was selected because the modes’ group velocity curves were relatively well separated and because their slopes over this region were significantly different.

Figure 7 shows the same arrival times as in Fig. 3, but

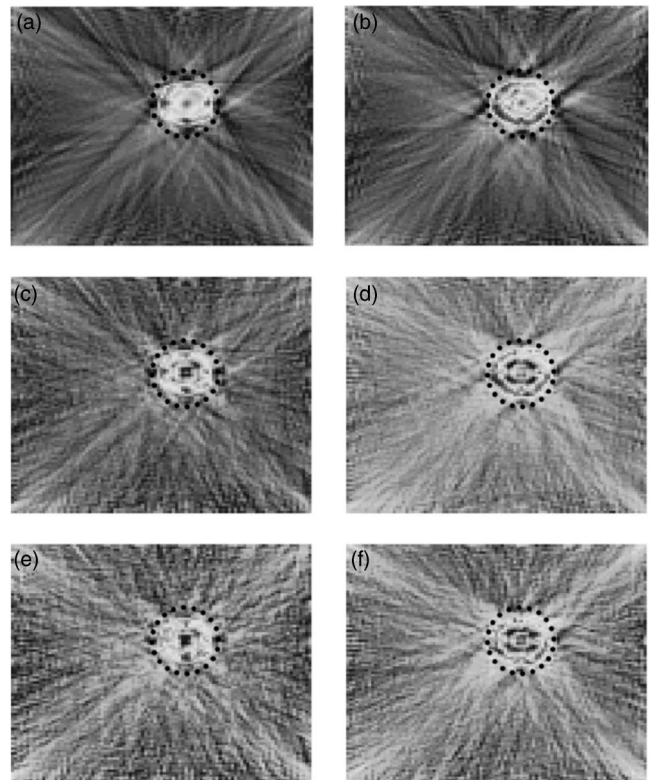


FIG. 11. Tomographic reconstructions for the first three modes of a 3.17-mm-thick plate sample with a 2.2-in.-diam flat-bottomed hole. Thickness loss within the flaw was 60%. (a) S_0 , unsorted. (b) S_0 , sorted. (c) A_0 , unsorted. (d) A_0 , sorted. (e) A_1 , unsorted. (f) A_1 , sorted. The dashed line in each reconstruction represents the actual location and size of the flaw.

separates them by mode and has the additional frequency walk information. The fastest mode, A_1 , is a smooth distribution between the different frequencies. The second peak’s arrival time can again be seen to jump from a time consistent with the second arriving mode, A_0 , to one consistent with the third arriving mode, S_0 , at about receiver position 61. It can then be seen to oscillate between the two values for a period.

Figure 8(a) shows the first 100 points of the frequency walk data for the first mode’s arrival time in the same flat-bottomed hole sample as above. Some frequencies interact with the flaw before the other frequencies. Also, due to either inconsistent coupling of the moving contact transducer or mode interference as described above, in the middle of the flawed area the first arriving peak in the higher frequencies can be seen to jump from the first mode to the second mode [see Fig. 8(b)]. Thus, even within the suspected flaw regions the data needs to be sorted so that the correct arrivals are included in the correct mode series.

The frequency walk data actually provides enough information to determine which mode series the arrival times should belong to. Figure 9(a) is a plot of the first three arrival times versus frequency number for pitch-catch position 65 (where the first peak’s arrival for the higher frequencies seems to jump to the second mode). From this graph it is seen that the first arrival for the higher frequencies, along with the second arrival for the lower frequencies, forms a line. This shows that something is occurring to mask the first

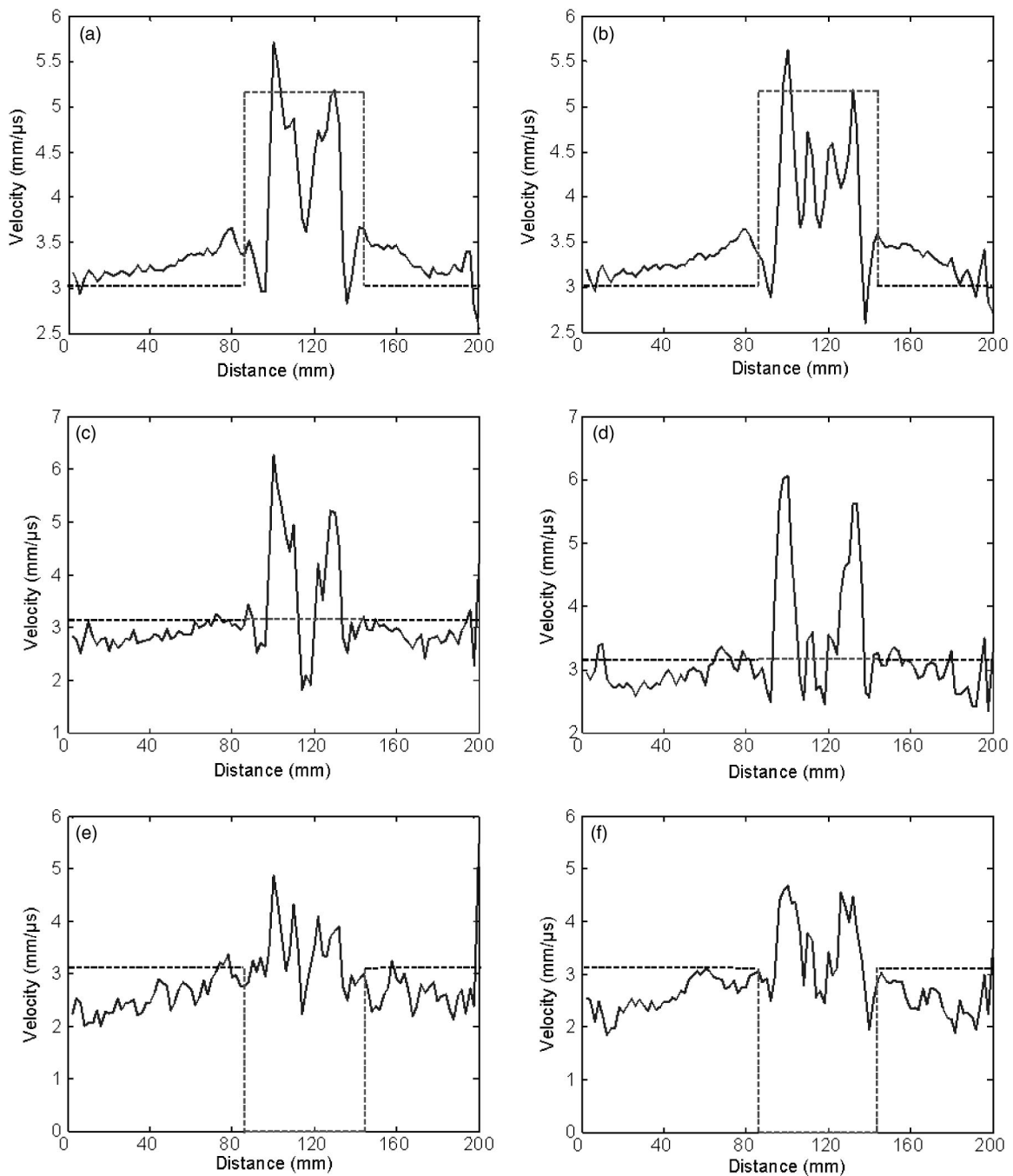


FIG. 12. Group velocity line plots for the first three modes of a vertical line through the center of the flat-bottomed hole. The dashed line represents the mode's theoretical group velocity for the measured frequency-thickness product outside and within the flaw. (a) S_0 , unsorted. (b) S_0 , sorted. (c) A_0 , unsorted. (d) A_0 , sorted. (e) A_1 , unsorted. (f) A_1 , sorted.

mode arrival and that the higher frequencies are actually locating the second mode as the first arriving peak. In the same way, Fig. 9(b) shows a similar plot for waveform 598 (the 98th receive position for the 5th transmitter position) in the flawed sample. In this graph, it can be seen that the first arrival for 900 kHz was missed. Therefore, the first arrival will have been marked as a potential flaw area by the sorting algorithm. However, the frequency data show that the first arrival actually belongs with the second mode and can be resorted without affecting the resulting reconstruction.

Therefore, using both the predicted mode curves and the frequency data, the multiple peak arrival times can be sorted

into the correct mode series both inside and outside a flawed region. First, each of the original peak arrival series is fitted with a third-order polynomial. Then each arrival time is compared with these expected mode arrivals. If the deviation between the actual arrival and the expected arrival is greater than $1.5 \mu\text{s}$, then the data point is marked as a potential flaw area. After all the arrivals of each frequency scan have been marked, the predicted mode values are updated by fitting only the arrivals which were not marked as a potential flaw. The nonflaw points are then sorted by the same algorithm used above to sort the clean plate arrivals. Finally, the flaw points are sorted using the frequency walk data. By sorting

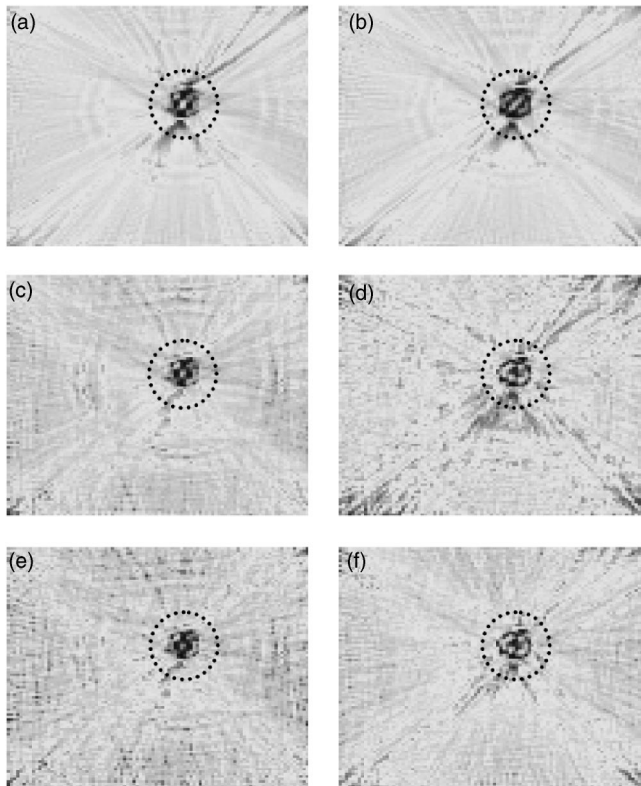


FIG. 13. Tomographic reconstructions for the first three modes of a 3.17-mm-thick plate sample with a 2.2-in.-diam, successively dished-out flat-bottomed hole. Maximum thickness loss within the center of the flaw was 60%. (a) S_0 , unsorted. (b) S_0 , sorted. (c) A_0 , unsorted. (d) A_0 , sorted. (e) A_1 , unsorted. (f) A_1 , sorted. The dashed line in each reconstruction represents the actual location and size of the flaw.

the frequency data into linear series as in Fig. 9, the arrival times within a flaw region can be sorted to the correct mode series.

Figure 10 shows the results of the sorting algorithm for the frequency walk data of the first 100 pitch-catch positions on the flat-bottomed hole sample. Figure 11 also shows the effects of the sorting algorithm on the tomographic reconstructions of the flat-bottomed hole sample for the first three modes. These reconstructions are for only the two parallel crosshole projections as opposed to the full six projections in multiple crosshole tomography so there are some unavoidable artifacts in the reconstructed images. It can be seen that the flaws are more accurately sized by the sorted data than the unsorted mode data. This is also illustrated in Fig. 12 where a velocity line plot is given for a vertical line that runs through the center of the flaw. The dashed lines in Fig. 12 represent the theoretical group velocities for the measured frequency-thickness (fd) product both within and outside of the flaw. The second arriving mode, A_0 , is not very dispersive and thus its theoretical group velocity does not change much within the flaw. However, the reconstructed data show a higher velocity within the flaw. Why this occurs has not been explored yet, but it is expected that this happens because of scattering, diffraction, and energy being converted from one mode to another as the Lamb wave modes interact with the flawed region. Similarly, the fd product within the flaw is also below the cutoff frequency of the A_1 mode.

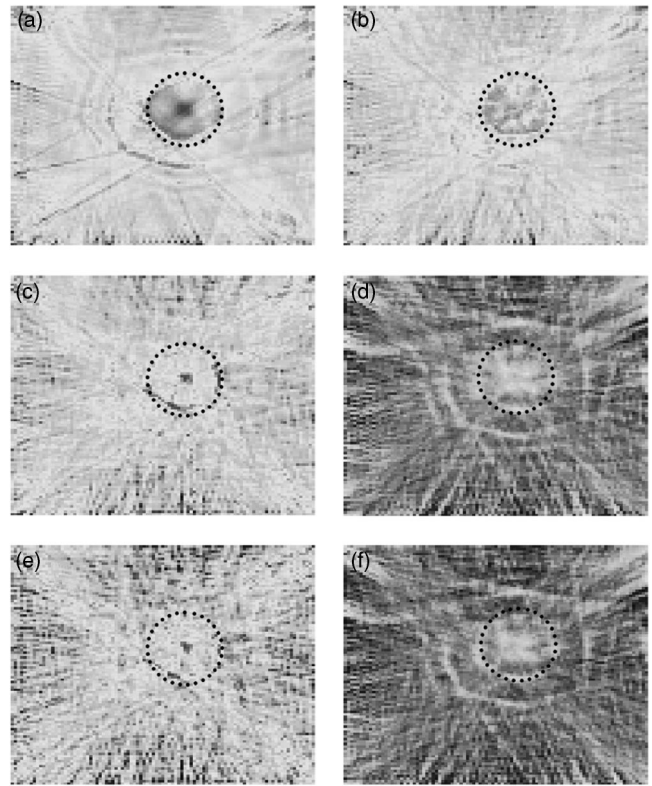


FIG. 14. Tomographic reconstructions for the first three modes of a 3.17-mm-thick plate sample with a 51-mm-diam circular flat-bottomed hole. Thickness loss was only 12% within the flaw. (a) S_0 , unsorted. (b) S_0 , sorted. (c) A_0 , unsorted. (d) A_0 , sorted. (e) A_1 , unsorted. (f) A_1 , sorted. The dashed line in each reconstruction represents the actual location and size of the flaw.

Again, the group velocity profile most likely shows a non-zero value because of the effects described above.

Similar results have been found for two other flawed samples. Figure 13 shows the reconstructions for the first three modes of a sample with a gradually dished out flat-bottomed hole the same diameter as the flat-bottomed hole sample above. The maximum thickness loss at the center of this flaw was 60%. The plate's thickness was again 3.17 mm. Finally, Fig. 14 shows a 2-in. flat-bottomed hole with a 12% thickness loss. It can be seen for this subtle flaw that the sorting algorithm improved the overall shape of the flaw for the S_0 mode.

The effects of the mode sorting algorithm on the tomographic reconstruction's quality can also be quantified with the following image quality parameters—contrast-to-noise ratio (CNR), flaw signal-to-noise ratio (FSNR), and speckle signal-to-noise ratio (SNR). These image quality parameters are similar to those defined by Behar *et al.*²⁶

$$\begin{aligned} \text{CNR} &= (\mu_s - \mu_n) / \sigma_n, \\ \text{FSNR} &= (\mu_s - \mu_n) / \sqrt{\sigma_s^2 + \sigma_n^2}, \\ \text{SNR} &= \mu_n / \sigma_n, \end{aligned} \quad (2)$$

where μ_s is the mean of the flaw data, μ_n is the mean of the noise outside of the flaw area, σ_n is the standard deviation of the noise outside of the flaw, and σ_s is the standard deviation of the flaw data. Table I summarizes their results. In order to

TABLE I. Image quality parameters for the four flawed samples (plate 1—flat bottomed hole; plate 2—successively dished out flat bottomed hole; plate 3—Thin flat bottomed hole, 2-in. diameter).

	Unsorted Plate 1—900 kHz			Sorted Plate 1—900 kHz			% Difference		
	S_0	A_0	A_1	S_0	A_0	A_1	S_0	A_0	A_1
CNR	2.56	1.80	1.36	2.85	2.36	1.85	11	31	36
FSNR	0.76	0.70	0.65	0.89	0.80	0.82	16	14	26
SNR	14.28	7.25	6.35	15.27	8.30	7.06	7	14	11
	Unsorted Plate 2—900 kHz			Sorted Plate 2—900 kHz			% Difference		
	S_0	A_0	A_1	S_0	A_0	A_1	S_0	A_0	A_1
CNR	-0.60	-0.34	-0.09	-0.72	-0.42	-0.10	19	23	4
FSNR	-0.24	-0.19	-0.06	-0.28	-0.21	-0.05	18	10	-11
SNR	10.35	7.51	6.29	10.05	9.06	8.57	-3	21	36
	Unsorted Plate 3—900 kHz			Sorted Plate 3—900 kHz			% Difference		
	S_0	A_0	A_1	S_0	A_0	A_1	S_0	A_0	A_1
CNR	1.44	0.84	0.57	1.20	0.84	0.77	-17	0	35
FSNR	1.13	0.64	0.44	0.92	0.65	0.59	-19	2	34
SNR	18.6	7.94	6.67	16.93	5.57	6.10	-9	-30	-9

accurately determine the flawed region from the unflawed region for the calculation of these parameters, the location and size of the flaws were carefully measured and the corresponding pixel values were used to select the correct areas in the images. It can be seen that the sorting algorithm does improve the overall image quality of the individual modes for each sample. From this data it can be seen that the sorting algorithm improves the contrast-to-noise ratio of the flaw for the second two modes the best. This is an important result, since it enhances the capability to automatically extract the flawed region from the unflawed region computationally.

V. CONCLUSION

This work, along with Hou,²¹ has laid the groundwork for the accurate extraction of multiple mode arrival times. In addition to platelike structures, it has also been shown by Leonard and Hinders²⁷ that pipelike objects can be treated as plates wrapped around upon themselves. Therefore, the same tomographic and arrival time sorting algorithms can also be used to inspect pipelike structures. The technique developed by Hou has the capability to extract more accurate arrivals and to ignore the unwanted pipe modes that exist in the more complicated pipe signals. Combined, the two techniques have the potential to lead to even better reconstructions that accurately locate and size flaws in both pipes and plates.

The sorting algorithm was demonstrated on four different plate samples. It was seen that after sorting the arrival times, the reconstructions did a better job of sizing the flaws. Another result from these tests was that after sorting, the tomographic data better represented the actual velocity data within the flawed area for the first mode and the deeper flaws showed up better in the reconstructions. In contrast, the sorting algorithm improved the quantitative accuracy—both in size and velocity reconstruction—for the subtle flaw, but the actual reconstructed images for the A_0 and A_1 modes did not highlight the flaw as well. However, this demonstrates the validity of the statement that the individual modes will interact with the flaws differently. The A_0 is nondispersive around the frequency-thickness product that we are operating and

thus we would not expect to see the flaw in these reconstructions unless other effects such as scattering, diffraction, or mode conversion are occurring.

Future work needs to pursue the application of the sorting and frequency compounding algorithms combined with the DWFP technique to pipelike structures. In order for the helical arrivals to be obtained accurately, a systematic study needs to be conducted on pipe data in order to obtain the correct fingerprint patterns for the different arrivals. Furthermore, as suggested by Hou,²¹ different wavelets need to be studied to determine which basis is best to solve this problem.

Finally, a systematic theoretical and experimental study of a representative flaw would be extremely beneficial. Since the calculations of the theoretical arrival times of the various modes in a flawed sample are nontrivial, it is hard to provide an accurate measure of the success of the multi-mode extraction algorithms. If a systematic study was done with a simple type of flaw that was slowly enlarged, both theoretical simulations—such as FEM, BEM, FDTD, etc.—and experimental data could be collected. This would allow for a comparison of the DWFP and sorting algorithms to theoretical data and would provide a clearer understanding of how the individual guided wave modes interact with a characteristic flaw.

¹D. P. Jansen and D. A. Hutchins, "Lamb Wave Tomography," IEEE 1990 Ultrasonics Symposium, 1990, Vol. 2, pp. 871–874.

²D. A. Hutchins, D. P. Jansen, and C. Edwards, "Lamb-wave tomography using non-contact transduction," Ultrasonics **31**(2), 97–103 (1993).

³D. P. Jansen and D. A. Hutchins, "Immersion tomography using Rayleigh and Lamb waves," Ultrasonics **30**(4), 245–254 (1992).

⁴D. P. Jansen, D. A. Hutchins, and J. T. Mottram, "Lamb wave tomography of advanced composite laminates containing damage," Ultrasonics **32**(2), 83–89 (1994).

⁵W. Wright, D. A. Hutchins, D. P. Jansen, and D. Schnidel, "Air-coupled Lamb wave tomography," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **44**(1), 53–59 (1997).

⁶Y. Nagata, J. D. Huang, J. D. Achenbach, and S. Krishnaswamy, "Lamb Wave Tomography Using Laser-Based Ultrasonics," Rev. Prog. Quant. Nondestr. Eval. **14**, 561–568 (1995).

⁷F. L. Degertekin, J. Pei, B. T. Khuri-Yakub, and K. C. Saraswat, "In situ

- acoustic temperature tomography of semiconductor wafers,” *Appl. Phys. Lett.* **64**(11), 1338–1340 (1994).
- ⁸J. C. P. McKeon, “Tomography Applied to Lamb Wave Contact Scanning,” Ph.D. dissertation Dept. of Applied Science, College of William and Mary, Williamsburg, VA, 1998.
- ⁹M. K. Hinders and J. C. P. McKeon, “Lamb Wave Tomography for Corrosion Mapping,” in *Proceedings of the 2nd Joint NASA/FAA/DoD Conference on Aging Aircraft*, 1999, pp. 732–740.
- ¹⁰M. K. Hinders, E. V. Malyarenko, and J. C. P. McKeon, “Contact scanning Lamb wave tomography,” *J. Acoust. Soc. Am.* **104**, 1790(A) (1998).
- ¹¹J. C. P. McKeon and M. K. Hinders, “Parallel projection and crosshole contact scanning Lamb wave tomography,” *J. Acoust. Soc. Am.* **106**, 2568–2577 (1999).
- ¹²J. C. P. McKeon and M. K. Hinders, “Lamb Wave Contact Scanning Tomography,” in *Review of Progress in QNDE* (Plenum, New York, 1999), pp. 951–958.
- ¹³B. P. Hildebrand, T. J. Davis, G. J. Posakony, and J. C. Spanner, “Lamb Wave Tomography for Imaging Erosion/Corrosion in Piping,” *Rev. Prog. Quant. Nondestr. Eval.* **18**, 967–973 (1999).
- ¹⁴E. V. Malyarenko and M. K. Hinders, “Fan beam and double crosshole Lamb wave tomography for mapping flaws in aging aircraft structures,” *J. Acoust. Soc. Am.* **108**, 1631–1639 (2000).
- ¹⁵E. V. Malyarenko, J. S. Heyman, and M. K. Hinders, “Lamb Wave Tomography for Monitoring Aircraft Structural Integrity,” in *USAF Aircraft Structural Integrity Program Conference Proceedings*, 2001.
- ¹⁶E. V. Malyarenko, “Lamb Wave Diffraction Tomography,” Ph.D. dissertation, Dept. of Applied Science, College of William and Mary, Williamsburg, VA, 2000.
- ¹⁷E. V. Malyarenko and M. K. Hinders, “Ultrasonic Lamb wave diffraction tomography,” *Ultrasonics* **39**(4), 269–281 (2001).
- ¹⁸E. V. Malyarenko and M. K. Hinders, “Comparison of Double Crosshole and Fanbeam Lamb Wave Tomography,” in *Review of Progress in QNDE* (Plenum, New York, 2001), pp. 732–739.
- ¹⁹M. K. Hinders, K. R. Leonard, and E. V. Malyarenko, “Blind test of Lamb wave diffraction tomography,” in *Review of Progress in QNDE* (AIP, Melville, NY, 2000), pp. 173–180.
- ²⁰K. R. Leonard, E. V. Malyarenko, and M. K. Hinders, “Ultrasonic Lamb wave tomography,” *Inverse Probl.* **18**, 1795–1808 (2002).
- ²¹J. Hou, “Ultrasonic signal detection and recognition using dynamic wavelet fingerprints,” Ph.D. dissertation, Dept. Applied Science, College of William and Mary, Williamsburg, VA, 2004.
- ²²J. Hou and M. K. Hinders, “Dynamic Wavelet Fingerprint Identification of Ultrasound Signals,” *Mater. Eval.* **60**(9), 1089–1093 (2002).
- ²³G. Strang, “Wavelets,” *Am. Sci.* **1994**, 250–255.
- ²⁴A. Abbate, J. Koay, J. Frankel, S. C. Schroeder, and P. Das, “Signal detection and noise suppression using a wavelet transform signal processor: application to ultrasonic flaw detection,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **44**, 14–26 (1997).
- ²⁵R. R. Coifman and D. L. Donoho, “Translation invariant de-noising,” *Lect. Notes Stat.* **103**, 125–150 (1995).
- ²⁶V. Behar, D. Adam, and Z. Friedman, “A new method of spatial compounding imaging,” *Ultrasonics* **41**(5), 377–384 (2003).
- ²⁷K. R. Leonard and M. K. Hinders, “Guided wave helical ultrasonic tomography of pipes,” *J. Acoust. Soc. Am.* **114**, 767–774 (2003).

Regularization method for measurement of structural intensity using nearfield acoustical holography

Kenji Saijyou^{a)} and Chiaki Okawara^{b)}

Fifth Research Center, Technical R & D Institute, Japan Defense Agency, 3-13-1 Nagase, Yokosuka, 239-0826, Japan

(Received 13 August 2003; revised 1 July 2005; accepted 31 January 2005)

The regularization method for measurement of structural intensity using nearfield acoustical holography is proposed. Spatial derivatives of normal displacement are necessary to obtain the structural intensity. The derivative operations amplify high-wave-number components of measurement noise. Therefore, the estimation of an appropriate wave-number filter is crucial for implementation of the measurement of structural intensity. In conventional methods, this wave-number filter is determined from the flexural wavelength. And the same wave-number filter is applied to obtain all spatial derivatives. As a result, structural intensity obtained from the pressure hologram, whose signal-to-noise ratio is low, is seriously contaminated by the noise. To overcome this difficulty, regularization theory is applied to determine the appropriate wave-number filter for each order of derivatives. The effectiveness of the proposed method is demonstrated by experiments. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1875652]

PACS numbers: 43.40.Dx, 43.40.Ey, 43.20.Ye [EGW]

Pages: 2039–2045

I. INTRODUCTION

Measurement of power flow in structures is essential for the vibration control and noise reduction. Many researchers have studied the measurement techniques of this structure-borne power flow, called structural intensity (SI). In the 1970s, Noiseux¹ and Pavic² introduced the formulation of the SI for thin beams and plates. To obtain the SI, spatial derivatives of the normal displacement of structure are necessary. This normal displacement is measured by various kinds of devices such as accelerometers,^{1,3–5} strain gauges,⁶ and laser Doppler vibrometer (LDV).^{7–10} Nearfield acoustical holography (NAH)¹¹ is also applied to obtain the normal displacement.^{12,13}

Strain gauges or accelerometers pick up the surface vibration directly to yield the local SI. However, we must attach many sensors on the structure if the SI is to be measured over the whole structure, and these attached sensors significantly affect the SI. On the contrary, noncontacting measurement methods based on LDV or NAH can yield the SI over the whole structure, and the method based on NAH enables us to obtain SI and acoustic intensity (AI) simultaneously.^{12,13} Therefore, NAH-based method is suitable to analyze the interrelationship between SI and AI.

Spatial derivatives of normal displacement are calculated by a finite difference (FD) approximation method^{1–7} or fast Fourier transform (FFT)-based method.^{8–10,12–16} The FD method is appropriate to measure local SI because of the simplicity of the FD method. However, the derivative operations of the FD method significantly magnify the noise of higher-order spatial derivatives. To avoid this noise amplification, the velocity or pressure data must be measured on a finely meshed contour.^{6,17} On the contrary, the FFT-based

method enables us to obtain spatial derivatives over the whole structure, and the amplification of noise is suppressed by applying a wave number-space (K-space) filter. In this method, the estimation of the appropriate filter is crucial. If the filter size is too large, the noise cannot be removed and the SI may be noisy. If the size is too small, useful information to locate a vibration source cannot be obtained.

In conventional FFT-based methods, the K-space filter is determined by a trial and error process,⁸ or determined from the flexural wavelength only.^{12,15,16} The same filter is used to evaluate all derivatives. Therefore, it is difficult to obtain SI by NAH from the low signal-to-noise ratio (SNR) measured pressure hologram using the K-space filter estimated by the conventional method.

In this paper, the estimation method of an appropriate k -space filter for each order of derivatives is proposed. The method utilizes the improved Tikhonov regularization filter¹⁸ (abbreviated as “Tikhonov filter” hereafter). The regularization parameter is determined by the modified generalized cross validation (GCV) method. As demonstration examples, experimental investigations are carried out to show the effectiveness of the proposed method.

II. STRUCTURAL INTENSITIES OF THIN PLATE

Before discussion of the regularization for measurement of SI, we review the basic formulas. First, let us describe the SI of thin plate below. If a plate is located on the x - y plane, the formulation for SI² is

$$I_x = D \left\langle \frac{\partial}{\partial x} (\nabla^2 w) \frac{\partial w}{\partial t} - \left(\frac{\partial^2 w}{\partial x^2} + \nu \frac{\partial^2 w}{\partial y^2} \right) \frac{\partial^2 w}{\partial x \partial t} - (1 - \sigma) \frac{\partial^2 w}{\partial x \partial y} \frac{\partial^2 w}{\partial x \partial t} \right\rangle_T$$

^{a)}Electronic mail: saiyou-kenji@jcom.home.ne.jp

^{b)}Electronic mail: okawara@jda-trdi.go.jp

$$I_y = D \left\langle \frac{\partial}{\partial y} (\nabla^2 w) \frac{\partial w}{\partial t} - \left(\frac{\partial^2 w}{\partial y^2} + \nu \frac{\partial^2 w}{\partial x^2} \right) \frac{\partial^2 w}{\partial y \partial t} - (1 - \sigma) \frac{\partial^2 w}{\partial x \partial y} \frac{\partial^2 w}{\partial y \partial t} \right\rangle_T, \quad (1)$$

where outer bracket $\langle \rangle_T$ indicates time average, E is the Young's modulus, σ is the Poisson's ratio, h is the plate thickness, w is the normal displacement, and $D = Eh^3/[12(1 - \sigma^2)]$. To calculate Eq. (1), spatial derivatives of the normal displacement are necessary. These spatial derivatives are easily obtained in K-space from the following relationship:^{8,12,15}

$$\frac{\partial^m}{\partial x^m} \frac{\partial^n}{\partial y^n} [w(x, y, 0)] = \mathcal{F}^{-1} \{ (ik_x)^m (ik_y)^n \mathcal{F}[w(x, y, 0)] \}, \quad (2)$$

where \mathcal{F} and \mathcal{F}^{-1} represent the two-dimensional forward and inverse Fourier transforms in x and y , and k_x and k_y are wave numbers in the x and y direction, respectively. The normal displacement of the plate is reconstructed by NAH¹¹ from the measured pressure, and SI of thin plate is calculated by Eqs. (1) and (2).

We have not considered the influence of measurement noise in these formulations of SI described in this section. In the practical NAH experiment, the reduction of the influence of measurement noise is crucial. In the next section, the regularization method for measurement of SI is explained in detail.

III. REGULARIZATION FOR SPATIAL DERIVATIVES

The Tikhonov filter for Cartesian coordinates is described as¹⁸

$$F^{\alpha,1}(k_x, k_y, \lambda_{k_x, k_y}) = \frac{|\lambda_{k_x, k_y}|^2}{|\lambda_{k_x, k_y}|^2 + \alpha [\alpha / (\alpha + |\lambda_{k_x, k_y}|^2)]^2}, \quad (3)$$

where

$$\lambda_{k_x, k_y} = \frac{\rho c k}{i \sqrt{k_x^2 + k_y^2 - k^2}} e^{-\sqrt{k_x^2 + k_y^2 - k^2} (z_{\text{mea}} - z_0)}. \quad (4)$$

is called the propagator,¹¹ ρ is the density of medium, c is the sound velocity, $k = 2\pi f/c$, f is the frequency, and $z = z_{\text{mea}}$ and $z = z_0$ are measurement and reconstruction planes, respectively. Let $\dot{w}(\mathbf{r}_0) = \dot{w}(x, y, z_0)$ be the normal velocity of the structure surface and $p(\mathbf{r}_{\text{mea}}) = p(x, y, z_{\text{mea}})$ be the hologram pressure on the measurement plane. The harmonic time dependence $e^{i\omega t}$ is suppressed throughout this paper. The following relationship exists between them:¹¹

$$w(\mathbf{r}_0) = \dot{w}(\mathbf{r}_0)/(i\omega) = \mathcal{F}^{-1} [\mathcal{F}\{p(\mathbf{r}_{\text{mea}})\} \lambda_{k_x, k_y}^{-1}] / (i\omega). \quad (5)$$

This is the reconstruction equation. Here, λ_{k_x, k_y}^{-1} is called the inverse propagator. To suppress the amplification of noise in high-wave-number region, the Tikhonov filter is applied. Thus Eq. (5) is modified as

$$\tilde{w}(\mathbf{r}_0) = \mathcal{F}^{-1} [\mathcal{F}\{p(\mathbf{r}_{\text{mea}})\} F^{\alpha,1}(k_x, k_y, \lambda_{k_x, k_y}) \lambda_{k_x, k_y}^{-1}] / (i\omega). \quad (6)$$

Here, tilde indicates the smoothed variable. To obtain spatial derivatives of the normal displacement from the measured pressure, Eq. (6) is substituted into Eq. (2) as

$$\begin{aligned} & \frac{\partial^m}{\partial x^m} \frac{\partial^n}{\partial y^n} [\tilde{w}(\mathbf{r}_0)] \\ &= \mathcal{F}^{-1} \{ \mathcal{F}\{p(\mathbf{r}_{\text{mea}})\} F^{\alpha,1}(k_x, k_y, \lambda_{k_x, k_y}) \\ & \quad \times \lambda_{k_x, k_y}^{-1} (ik_x)^m (ik_y)^n \} / (i\omega). \end{aligned} \quad (7)$$

Equation (7) is a typical FFT-based method. Here, the filter $F^{\alpha,1}(k_x, k_y, \lambda_{k_x, k_y})$ is appropriate to obtain the normal velocity $\dot{w}(\mathbf{r}_0)$ because the inverse propagator λ_{k_x, k_y}^{-1} , which describes the relationship between the measured pressure $p(\mathbf{r}_{\text{mea}})$ and the normal velocity $\dot{w}(\mathbf{r}_0)$, is treated as the eigenvalue of the Tikhonov filter. However, this filter is not appropriate to obtain the spatial derivative of the normal displacement from the measured pressure, because $(ik_x)^m (ik_y)^n$ is multiplied to the K-space representation of the measured pressure $\mathcal{F}\{p(\mathbf{r}_{\text{mea}})\}$ and this filter does not sufficiently suppress the growth of the high-wave-number noise which contaminates the reconstructed result.

To suppress the amplification of high-wave-number noise, Morikawa¹⁵ proposed that the circ function should be applied for the k -space filter to the normal velocity as

$$\begin{aligned} & \frac{\partial^m}{\partial x^m} \frac{\partial^n}{\partial y^n} [\tilde{w}(\mathbf{r}_0)] \\ &= \mathcal{F}^{-1} \{ \mathcal{F}\{p(\mathbf{r}_{\text{mea}})\} \text{circ}(k_x, k_y, k_{\text{cut}}) \\ & \quad \times F^{\alpha,1}(k_x, k_y, \lambda_{k_x, k_y}) \lambda_{k_x, k_y}^{-1} (ik_x)^m (ik_y)^n \} / (i\omega), \end{aligned} \quad (8)$$

$$\text{circ}(k_x, k_y, k_{\text{cut}}) = \begin{cases} 1, & k_x^2 + k_y^2 < k_{\text{cut}}^2 \\ 0, & k_x^2 + k_y^2 \geq k_{\text{cut}}^2. \end{cases} \quad (9)$$

Here, the cutoff wave number satisfies the relationship as $1.5k_{\text{flex}} < k_{\text{cut}} < 1.8k_{\text{flex}}$, $k_{\text{flex}} (= [h\rho_0(2\pi f)^2/D]^{1/4})$ is the flexural wave number, and ρ_0 is the density of the plate. Nejade¹⁶ proposed that a band-pass filter should be applied for the K-space filter as the normal velocity as

$$\begin{aligned} & \frac{\partial^m}{\partial x^m} \frac{\partial^n}{\partial y^n} [\tilde{w}(\mathbf{r}_0)] \\ &= \mathcal{F}^{-1} \left\{ \mathcal{F}\{p(\mathbf{r}_{\text{mea}})\} \left[\text{circ} \left(k_x, k_y, k_{\text{flex}} + \frac{2\pi}{L} \right) \right. \right. \\ & \quad \left. \left. - \text{circ} \left(k_x, k_y, k_{\text{flex}} - \frac{2\pi}{L} \right) \right] \right. \\ & \quad \left. \times F^{\alpha,1}(k_x, k_y, \lambda_{k_x, k_y}) \lambda_{k_x, k_y}^{-1} (ik_x)^m (ik_y)^n \right\} / (i\omega). \end{aligned} \quad (10)$$

Here L is the width of the measurement aperture. However, if the cutoff wave number is too small, the useful information

cannot be obtained to calculate SI, because the resolution of the reconstructed result is degraded. The amplification by the lower-order spatial derivative is relatively small; therefore, the lighter filtration is necessary to obtain the useful information when a lower-order derivative is reconstructed. When the cutoff wave number is too large, the amplification of noise in the high-wave-number region seriously contaminates the reconstructed result. The amplification by the higher-order derivative is relatively large; therefore, the deeper filtration is necessary to suppress the amplification of noise when a higher-order derivative is reconstructed. In other words, the appropriate cutoff wave number for circ function filter and band-pass filter are different at each frequency for each derivative order, and the appropriate cutoff wave number is not uniquely determined by conventional methods. Therefore, conventional methods are not appropriate for obtaining the SI when the SNR of the measured pressure hologram is low.

Now, we introduce the modification of the Tikhonov filter for spatial derivatives. The term, which is contained in Eq. (7),

$$\lambda_{k_x, k_y, (x, m), (y, n)}^{-1} = \lambda_{k_x, k_y}^{-1} (ik_x)^m (ik_y)^n, \quad (11)$$

expresses the relationship between the measured pressure and the spatial derivative of the normal displacement. Then, we can treat Eq. (11) as an eigenvalue of the Tikhonov filter, and the effect of the spatial derivative can be included in the determination process of regularization filter. The Tikhonov filter is modified for the spatial derivative as

$$F^{\alpha, 1}(k_x, k_y, \lambda_{k_x, k_y, (x, m), (y, n)}) = \frac{|\lambda_{k_x, k_y, (x, m), (y, n)}|^2}{|\lambda_{k_x, k_y, (x, m), (y, n)}|^2 + \alpha[\alpha + |\lambda_{k_x, k_y, (x, m), (y, n)}|^2]^2}, \quad (12)$$

and the spatial derivative is obtained as

$$\begin{aligned} & \frac{\partial^m}{\partial x^m} \frac{\partial^n}{\partial y^n} [\tilde{w}(\mathbf{r}_0)] \\ &= \mathcal{F}^{-1} \{ \mathcal{F} \{ p(\mathbf{r}_{\text{mea}}) \} F^{\alpha, 1}(k_x, k_y, \lambda_{k_x, k_y, (x, m), (y, n)}) \\ & \quad \times \lambda_{k_x, k_y, (x, m), (y, n)}^{-1} \} / (i\omega). \end{aligned} \quad (13)$$

Equation (12) is a Tikhonov filter for the spatial derivative. This filter is optimized for the order of derivative and SNR of the measured pressure.

IV. MODEL EXPERIMENT

In this section, the SI in a thin plate is measured to demonstrate the effectiveness of the proposed method. In experiments, natural frequencies at which the driving point is easily localized and other frequencies at which the localization of the driving point is difficult (this frequency is called as “non-natural frequency” hereafter) are selected. An automated scanning system was developed for measurement of airborne sound. An experimental model is SUS 304 stainless steel (the side is 500 mm × 600 mm × 5 mm). A shaker (Wil-

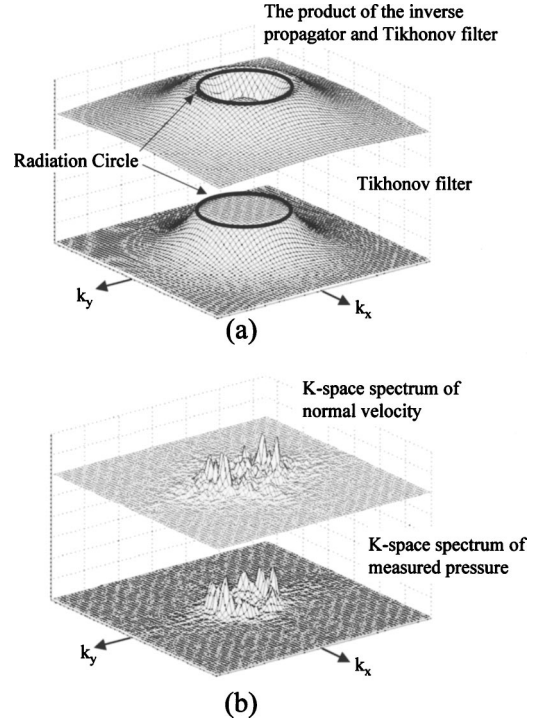


FIG. 1. K-space filter and K-space spectra. The frequency is 1144 Hz. (a) Tikhonov filter and the product of the inverse propagator and Tikhonov filter. (b) K-space spectra of measured pressure and reconstructed normal velocity.

coxon F3/Z3/F9) that vibrates the sample plate is attached at the center of the plate. The shaker is driven by a (300–6000 Hz) linear frequency modulation (LFM) signal. A Brüel & Kjaer Type 4182 microphone is placed at the prescribed position with scanners. Stepwise movement of the plate thus scans a two-dimensional measurement plane. Measured data of the measurement aperture are taken in 20-mm steps along x - and y -axis scanning. Thus, 64 × 64 points data make one hologram. The distance between the sample plate and the measurement hologram is 3 mm. The received and amplified signal from the microphone is digitized and transferred to a personal computer. Sampling inception time is accurately controlled by the computer, which acts as the reference source for holographic interference. The ends of the plate are sealed with rubber packing to approximate the simply supported boundary condition.

Figure 1 shows the K-space filter, K-space spectra of measured pressure and normal velocity. The plate is vibrated at the natural frequency of 1144 Hz at which the driving mobility shows a peak. In this paper, all data are normalized by the maximum value. The filter is the Tikhonov filter for normal velocity [Eq. (3)]. In Fig. 1(a), circles indicate radiation circles. The product of the inverse propagator and the Tikhonov filter show that (1) the amplitude inside the radiation circle is small, (2) the peak exists near the radiation circle, and (3) the amplitude outside the radiation circle is small. The K-space spectrum of the normal velocity is expressed as

$$\mathcal{F}[\tilde{w}(\mathbf{r}_0)] = \mathcal{F}[p(\mathbf{r}_{\text{mea}})] F^{\alpha, 1}(k_x, k_y, \lambda_{k_x, k_y}) \lambda_{k_x, k_y}^{-1}. \quad (14)$$

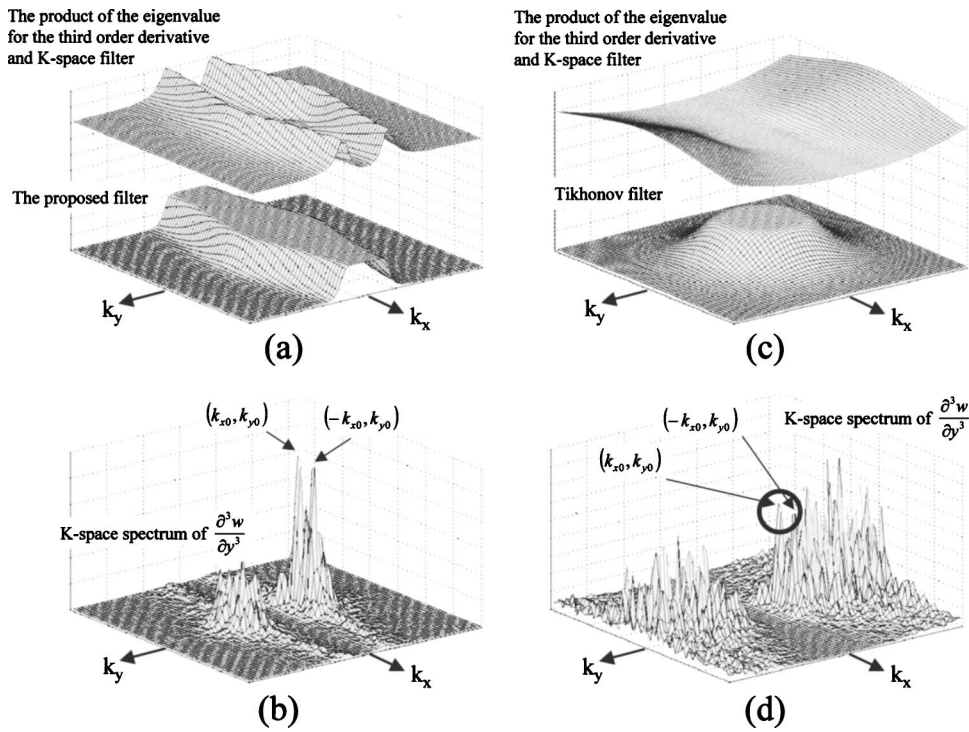


FIG. 2. K-space filter and K-space spectra. The frequency is 1144 Hz. (a) The proposed filter and the product of the eigenvalue for the third-order derivative $\lambda_{k_x, k_y, (x,0), (y,0)}^{-1}$ and the proposed filter. (b) K-space spectrum of $\partial^3 w / \partial y^3$ obtained by the proposed filter. (c) The Tikhonov filter for normal velocity and the product of eigenvalue for the third order derivative $\lambda_{k_x, k_y, (x,0), (y,0)}^{-1}$ and the Tikhonov filter. (d) K-space spectrum of $\partial^3 w / \partial y^3$ obtained by the Tikhonov filter. The circle shows local peaks that correspond to the peak of (b).

Therefore, the K-space spectrum of the normal velocity is wider than that of the measured pressure.

Figure 2 shows the K-space filter and K-space spectrum of the third-order spatial derivative of normal displacement. The frequency is 1144 Hz. Figure 2(a) shows the proposed filter for the third-order derivative and the product of the eigenvalue of the filter for the third-order derivative $\lambda_{k_x, k_y, (x,0), (y,3)}^{-1}$ [i.e., the product of the inverse propagator and $(ik_y)^3$] and the filter. The pass-band of the filter exists along the k_x axis, and the amplitude of the product of the eigenvalue and the filter is large near the k_x axis. Here, high-wave-number component is exponentially amplified in the NAH reconstruction process because of the multiplication of the inverse propagator. However, in this case, the amplification by the spatial derivative surpasses that by multiplication of the inverse propagator. In fact, the cube of the highest wave number of the y direction is $(2\pi/0.04)^3 \approx 3\,876\,000$, and the maximum amplitude of the inverse propagator is $e^{(2\pi/0.04) \cdot 0.005} \approx 345$. Therefore, the suppression of the growth of the noise due to the spatial derivative by the Tikhonov filter is appropriate to obtain the spatial derivative by Eq. (13). As the result, two peaks of the K-space spectrum of $\partial^3 w / \partial y^3$ [Fig. 2(b)] exist at $(\pm k_{x0}, k_{y0})$ ($\sqrt{k_{x0}^2 + k_{y0}^2} = k_{\text{flex}}$), and the amplitude $\partial^3 w / \partial y^3$ of the spectrum diminishes at the region which is far from the k_x axis. This result exhibits that the amplification of noise at the high-wave-number region is sufficiently suppressed by the proposed filter.

Figure 2(c) shows the Tikhonov filter and the product of the eigenvalue for third-order derivative $\lambda_{k_x, k_y, (x,0), (y,3)}^{-1}$ and the K-space filter. This Tikhonov filter is the same as that of Fig. 1(a). The Tikhonov filter cannot suppress the amplitude of the eigenvalue for third-order spatial derivatives, and the amplitude of the product of the propagator and the filter is

large at the region which is far from the k_x axis. As the result, the K-space spectrum of $\partial^3 w / \partial y^3$ [Fig. 2(d)] is seriously contaminated by noise. Two peaks in Fig. 2(d) marked with a circle corresponding to peaks in Fig. 2(b) are about to be hidden in some noise. Therefore, the Tikhonov filter for the normal velocity is not an appropriate filter for higher-order derivatives. To suppress this noise contamination, Morikawa¹⁵ and Nejade¹⁶ proposed the K-space filter whose cutoff wave number is lower. However, the size of the filter is sometimes too small, and useful information to calculate the SI may not be obtained.

To demonstrate the influence of K-space filters, the injected power calculated from the SI of the plate obtained by each K-space filter is compared. The injected power is given by¹²

$$\Pi_C = \oint_C \mathbf{I} \cdot \hat{n} \, dl, \quad (15)$$

Here, C is the contour surrounding each point, dl is the element of a square contour C , \mathbf{I} is the SI vector, and \hat{n} is the outward normal to the element of contour dl . Figure 3 shows the injected power at 1144 Hz. This frequency is the natural frequency. The SNR of the measured pressure hologram is 40 dB. The driving point is shown by a small triangle symbol, hereafter. Gray-scale plots show injected power and AI, white is positive, and black is negative, respectively. In this figure, the horizontal and vertical axes indicate x and y axes, respectively. Figure 3(a) represents the injected power reconstructed by the proposed method. The peak of the injected power corresponds to the driving point and comparison between the injected power and the AI [cf. Fig. 3(e)] shows fairly good matching at the periphery of the plate because Π_C is equal to the negative of the normal AI away from the driver.¹⁹ Figures 3(b) and (c) are the injected power recon-

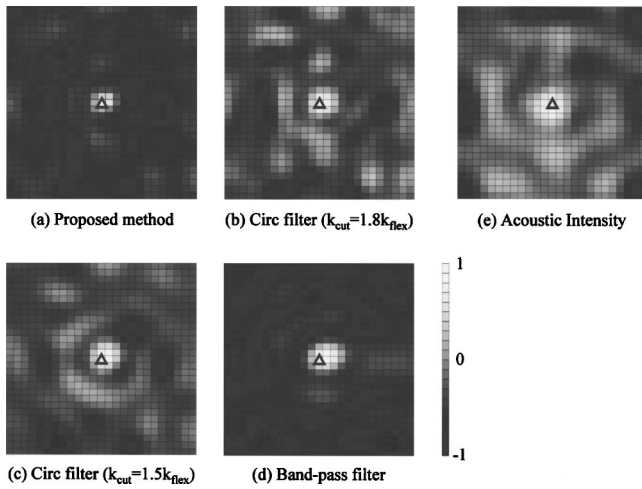


FIG. 3. The power injected to the plate. The frequency is 1144 Hz. SNR of the measured pressure is 40 dB. The triangle indicates the driving point. (a) The injected power by the proposed method. (b) The injected power by the circ function filter. The cutoff wave number is $1.8k_{flex}$. (c) The injected power by the circ function filter. The cutoff wave number is $1.5k_{flex}$. (d) The injected power by the band-pass filter. The center wave number is k_{flex} . (e) The normal component of the acoustic intensity radiated from the plate.

structed by circ function filter, which is proposed by Morikawa¹⁵ [cf. Eq. (8)], and Fig. 3(d) is that reconstructed by band-pass filter, which is proposed by Nejade¹⁶ [cf. Eq. (10)]. In Figs. 3(b)–(d), the peak of the injected power corresponds to the driving point. Comparison between the injected power and the AI shows fairly good matching away from the driver in Figs. 3(b) and (c). However, the region surrounding the driving point is contaminated by the amplification of noise in Figs. 3(b) and (c), because the cutoff filters are not selected appropriately. In Fig. 3(d), the injected power is not matched by the AI away from the driver, because the filter removes useful information.

Figure 4 shows the injected power at 1144 Hz. Random Gaussian noise was added to the measured pressure hologram with a 15-dB SNR. The (AI [Fig. 4(e)] is the same as Fig. 3(e) for comparison between reconstructed results of injected power and AI.) Differences between Figs. 3(b) and 4(b), Figs. 3(c) and 4(c), and Figs. 3(d) and 4(d) are relatively small, because the cutoff wave number is small and the influence of the noise is small. Comparison between Figs. 3(a) and 4(a) shows relatively large difference, especially in the vicinity of the driving point. However, the peak of the injected power corresponds to the driving point, and comparison between the injected power and the AI shows fairly good matching away from the driver. As a result, all methods considered here enable us to localize the driving point at 1144 Hz, because 1144 Hz is the natural frequency of the plate and the amplitude of the mode component in K-space surpasses noise components. However, the injected power reconstructed by the circ function filter shows that the region surrounding the driving point is contaminated by the amplification of noise and the injected power reconstructed by the band-pass filter is not matched by the AI away from the driver. On the contrary, the injected power reconstructed by the proposed filter shows fairly good matching with AI away from the driver.

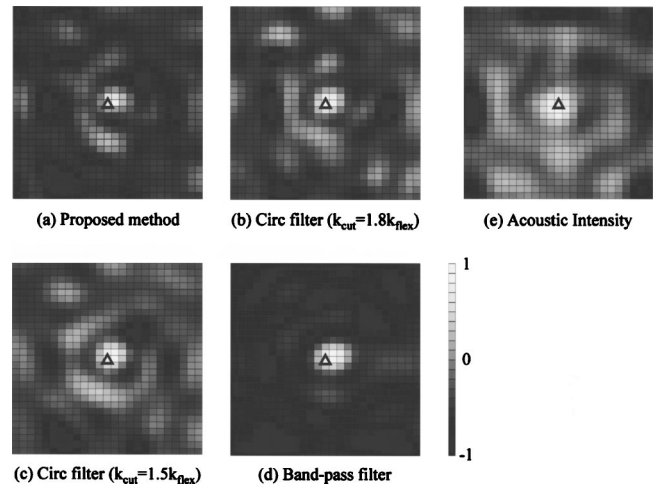


FIG. 4. The power injected to the plate. The frequency is 1144 Hz. SNR of the measured pressure is 15 dB. The triangle indicates the driving point. (a) The injected power by the proposed method. (b) The injected power by the circ function filter. The cutoff wave number is $1.8k_{flex}$. (c) The injected power by the circ function filter. The cutoff wave number is $1.5k_{flex}$. (d) The injected power by the band-pass filter. The center wave number is k_{flex} . (e) The normal component of the acoustic intensity radiated from the plate, which is reconstructed from the measured pressure without additional random Gaussian noise.

Figure 5 shows the injected power at 1923 Hz that is the non-natural frequency. The SNR of the measured pressure hologram is 40 dB. Figure 5(a) shows that the peak corresponds to the driving point and local peaks at the periphery of the plate are matched by the AI [cf. Fig. 5(e)]. Figures 5(b) and (c) show that the peak is exhibited at the upper-right corner of the plate and does not correspond to the driving point. The injected power at the driving point is negative. Figure 5(d) shows that the peak corresponds to the driving point; however, false local peaks are exhibited near the driving point and the injected power is not matched by the AI away from the driving point. As the result, the appropriate

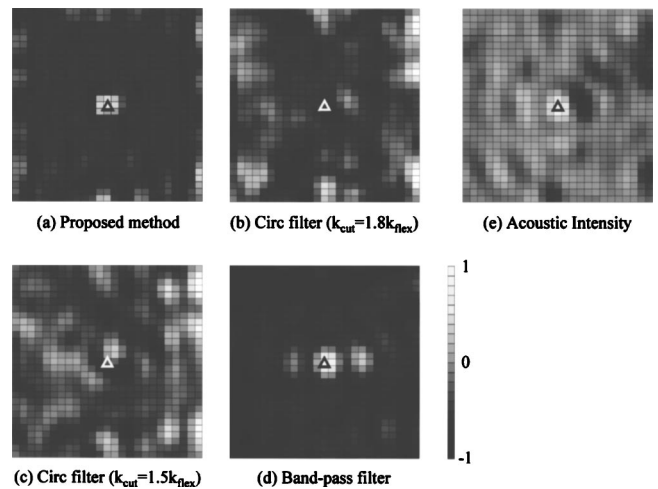


FIG. 5. The power injected to the plate. The frequency is 1923 Hz. SNR of the measured pressure is 40 dB. The triangle indicates the driving point. (a) The injected power by the proposed method. (b) The injected power by the circ function filter. The cutoff wave number is $1.8k_{flex}$. (c) The injected power by the circ function filter. The cutoff wave number is $1.5k_{flex}$. (d) The injected power by the band-pass filter. The center wave number is k_{flex} . (e) The normal component of the acoustic intensity radiated from the plate.

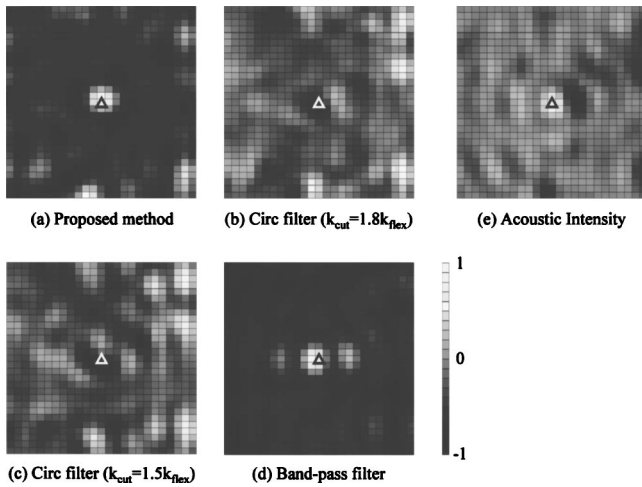


FIG. 6. The power injected to the plate. The frequency is 1923 Hz. SNR of the measured pressure is 15 dB. The triangle indicates the driving point. (a) The injected power by the proposed method. (b) The injected power by the circ function filter. The cutoff wave number is $1.8k_{flex}$. (c) The injected power by the circ function filter. The cutoff wave number is $1.5k_{flex}$. (d) The injected power by the band-pass filter. The center wave number is k_{flex} . (e) The normal component of the acoustic intensity radiated from the plate, which is reconstructed from the measured pressure without additional random Gaussian noise.

filter can only be estimated by the proposed method.

Figure 6 shows the injected power at 1923 Hz. Random Gaussian noise was added to the measured pressure hologram with a 15-dB SNR except for the AI [Fig. 6(e)]. In Fig. 6(a), the injected power shows a relatively large difference because of the noise influence. However, the peak corresponds to the driving point, and the comparison between Figs. 6(a) and (e) shows fairly good matching at the periphery of the plate. Figures 6(b) and (c) show that the peak does not correspond to the driving point. Comparison between Figs. 5(d) and 6(d) shows that the difference is relatively small, because the cutoff wave number is low. The peak corresponds to the driving point; however, false local peaks are exhibited near the driving point and the injected power is not matched by the AI away from the driving point. Here 1923 Hz is the non-natural frequency, therefore the vibration mode is not clearly shown, and the influence of the noise inside the circ filter is serious, especially in the higher-order spatial derivative [cf. Figs. 6(b) and (c)]. And the lack of the high-wavenumber component induces the false local peaks near the driving point [cf. Fig. 6(d)]. Therefore, conventional methods cannot bring us the appropriate filtration in this critical situation. On the contrary, the proposed method can localize the driving point without any false local peaks [cf. Fig. 6(a)].

V. CONCLUSIONS

A new regularization method for measurement of SI using NAH is presented. Since the K-space filter is determined from the flexural wavelength by conventional methods, it is difficult to measure SI by NAH from the low-SNR pressure hologram with the conventional filter. Therefore, the SI has not been applied to detect the vibration source in the noisy field, because the SNR of the measured pressure hologram is

low. To overcome this situation, the Tikhonov regularization filter is modified for each order of derivative. The regularization parameter for each derivative order is determined by modified GCV method using the measured pressure. Therefore, the K-space filter is optimized for low-SNR measured pressure.

The effectiveness of the proposed method is demonstrated by experiments. For comparison, the K-space filters estimated by the conventional methods are used to measure the SI. The experimental results of the SI in a thin plate are as follows:

- (1) The circ function filter proposed by Morikawa¹⁵ can localize the driving point at the natural frequency; however, the region surrounding the driving point is contaminated by the amplification of noise.
- (2) The circ function filter cannot localize the driving point at the non-natural frequency, and many false local peaks are exhibited.
- (3) The band-pass filter proposed by Najede¹⁶ can localize the driving point whether the frequency is the natural frequency or not; however, the injected power is not matched by the AI away from the driver at the natural frequency, and false local peaks are exhibited at the non-natural frequency.
- (4) The influences of noise to the reconstructed result by the conventional filter (i.e., circ function filter and band pass filter) are serious at the non-natural frequency.
- (5) On the contrary, the proposed filter can localize the driving point without any false local peaks whether the frequency is the natural frequency or not, and the influence of the noise to the reconstructed result is relatively small.

As a result, the availability of the proposed method is confirmed. This method may thus be recommended to apply the detection of the vibration source in an actual noisy field.

¹D. U. Noiseux, "Measurement of power flow in uniform beams and plate," *J. Acoust. Soc. Am.* **47**, 238–247 (1970).

²G. Pavic, "Measurement of structure borne wave intensity, Part I: Formulation of the methods," *J. Sound Vib.* **49**, 221–230 (1976).

³H. Zhou, N. Kojima, and M. Kato, "Vibration intensity measurement using three accelerometers," *Trans. Jpn. Soc. Mech. Eng.* **58**, 455–451 (1993).

⁴N. Tanaka, S. D. Snyder, Y. Kikushima, and M. Kuroda, "Vortex structural power flow in a thin plate and the influence on the acoustic field," *J. Acoust. Soc. Am.* **96**, 1563–1574 (1993).

⁵N. Tanaka, S. D. Snyder, and C. H. Hansen, "Velocity characteristics of the vibration intensity field in an actively controlled thin plate," *J. Acoust. Soc. Am.* **99**, 942–953 (1996).

⁶G. Pavic, "Measurement of vibration by strain gauges, Part I. Theoretical basis," *J. Sound Vib.* **102**, 153–163 (1985).

⁷S. I. Hayek, M. J. Pechersky, and B. C. Suen, "Measurement and analysis of near and far field structural intensity by scanning laser vibrometry," *Proceedings of the 3rd International Congress on Intensity Techniques*, Senlis, France (1990), pp. 281–288.

⁸Y. Zhang and J. A. Mann III, "Measuring the structural intensity and force distribution in plate," *J. Acoust. Soc. Am.* **99**, 345–353 (1996).

⁹Y. Zhang and J. A. Mann III, "Examples of using structural intensity and force distribution to study vibrating plates," *J. Acoust. Soc. Am.* **99**, 354–361 (1996).

¹⁰J. Pascal, T. Loyau, and X. Carniel, "Complete determination of structural intensity in plate using laser vibrometers," *J. Sound Vib.* **161**, 527–531 (1993).

- ¹¹E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic, London, UK, 1999).
- ¹²E. G. Williams, H. D. Dardy, and R. G. Fisk, "A technique for measurement of structure-borne intensity in plates," *J. Acoust. Soc. Am.* **78**, 2061–2068 (1985).
- ¹³J. C. Pascal, T. Loyau, and J. A. Mann III, "Structural intensity from spatial Fourier transformation and BAHIM acoustic holography method," *Proceedings of the 3rd International Congress on Intensity Techniques*, Senlis, France (1990), pp. 197–204.
- ¹⁴A. B. Spalding and J. A. Mann III, "Placing small constrained layer damping patches on a plate to attain global or local velocity changes," *J. Acoust. Soc. Am.* **97**, 3617–3624 (1985).
- ¹⁵R. Morikawa, S. Ueha, and K. Nakamura, "Error evaluation of the structural intensity measured with a scanning laser Doppler vibrometer and a k-space signal processing," *J. Acoust. Soc. Am.* **99**, 2913–2921 (1996).
- ¹⁶A. Nejade and R. Singh, "Flexural intensity measurement on finite plates using modal spectrum ideal filtering," *J. Sound Vib.* **256**, 33–63 (2002).
- ¹⁷C. B. Lee, Y. Iwaya, Y. Suzuki, and T. Sone, "A study on the measurement of one-dimensional vibration intensity carried by bending vibration in a beam," *J. Acoust. Soc. Jpn. (E)* **14**, 11–18 (1993).
- ¹⁸E. G. Williams, "Regularization method for near-field acoustic holography," *J. Acoust. Soc. Am.* **110**, 1976–1988 (2001).
- ¹⁹E. G. Williams, "Structural intensity in thin cylindrical shells," *J. Acoust. Soc. Am.* **89**, 1615–1622 (1991).

Smart panels with velocity feedback control systems using triangularly shaped strain actuators

Paolo Gardonio and Stephen J. Elliott

Institute of Sound and Vibration Research, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom

(Received 28 July 2004; revised 2 December 2004; accepted 25 December 2004)

In this paper we present a theoretical study on the active structural acoustic control of a new smart panel with sixteen triangularly shaped piezoelectric patch actuators, having their base edges evenly distributed along the perimeter of the panel, and velocity sensors positioned at the vertices opposite the base edges. The performance is assessed and contrasted with that of a conventional smart panel using a 4×4 array of square piezoelectric patch actuators evenly distributed over the surface of the panel with velocity sensors at their centers. For both systems the control effectiveness and stability of MIMO decentralized or SISO direct velocity feedback control architectures have been analyzed. The two control systems are arranged to generate active damping which reduces the response and sound radiation of the panel in the lightly damped and well separated low-frequency resonances. In particular the new control system can be seen as a set of sixteen “active wedges” which absorb energy from the incident flexural waves to the borders of the panel so that the panel could be considered anechoic. This study shows that the new arrangement with triangularly shaped actuators can achieve better control than the corresponding system using square actuators. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1863092]

PACS numbers: 43.40.Vn, 43.50.Ki [KAC]

Pages: 2046–2064

I. INTRODUCTION

In general, the low-frequency broadband sound transmission through a lightly damped and lightweight panel, such that its response is characterized by well separated resonances below the critical frequency, is controlled by the response of the panel itself at the resonance frequencies and by the radiation efficiencies of the resonant modes.¹ The sensitivity to the radiation efficiencies arises from both the acoustical excitation and sound radiation mechanisms.¹ In order to control low-frequency sound transmission through a panel it is therefore necessary to either move up its first few resonance frequencies by stiffening the panel or to apply damping treatments to the panel in such a way as to reduce its response at resonance frequencies.¹ Both types of treatments require substantial variations to the structure of the panel which have several drawbacks such as, for example, the change of geometry and weight of the partition and increase of costs. Alternatively, active control techniques could be used, the most attractive of which have the actuators and sensors integrated onto the partition itself in such a way as to create a smart panel for the implementation of active structural acoustic control (ASAC).^{2–5} The control system linking these actuators to these sensors increasingly uses a feedback rather than feed-forward arrangement, because of its ability to deal with broadband random or transient disturbances without an external reference signal.⁶

During the past decade a lot of research work has been carried out to develop ASAC systems to be embedded on thin partitions in order to form compact and possibly lightweight smart panels.^{2–5} Several configurations have been studied that could be grouped into two main families: first, single input single output (SISO) classic feedback control systems^{3,7} using collocated and dual^{8,9} distributed sensor-

actuator pairs which are arranged in such a way as to control radiation modes of panels^{2,10–12} and, second, SISO or multi-input multi-output (MIMO) modern feedback control systems^{3,7} where H_2 or H_∞ optimum state regulators are designed in a LQR/LQG framework, with a state observer based on radiation filters.^{13–15} The advantage of the classic approach is given by the possibility of implementing relatively simple direct feedback control loops that produce active damping⁷ and thus reduces the response of the panel at the low resonance frequencies without sound radiation spillover effects.¹⁶ However, several problems have been encountered in the development of truly collocated and dual^{8,9} distributed sensor-actuator pairs which would guarantee unconditionally stable feedback control loops.^{7,17–24} In contrast the modern approach uses much simpler arrays of sensors and actuators but rather complicated control systems with a state observer that inherently limits the control effectiveness as well as the robustness of the controller.^{3,20,24–37} In summary the development of SISO classic feedback control systems is held back by the difficulties encountered to develop collocated and dual distributed sensor-actuator pairs whose response functions have real part positive definite so that unconditionally stable direct feedback loops could be implemented.^{3,7} In contrast the development of SISO or MIMO modern feedback control systems is held back by the necessity of developing complicated state observers which should be robust to changes of the response of the panel due, for example, to variations of temperature, pressure loading, tensioning effects, etc.

In this context Petitjean and Legrain³⁸ have considered the possibility of simplifying the architecture of a MIMO feedback controller for a smart panel with a 5×3 array of closely located piezoelectric patch sensor-actuator pairs by

implementing decentralized feedback control loops with fixed control gains. In particular they compared the control effectiveness of the decentralized control system with that of a fully coupled MIMO feedback controller and found comparable control results for the two systems. Like a number of authors Elliott *et al.*³⁹ have shown that, provided the sensors and actuators are arranged in collocated and dual pairs, unconditionally stable decentralized feedback loops with fixed gains could be implemented. They have also shown that, if the control systems are arranged to implement active damping, for example with direct velocity feedback loops, then as the control gains are raised up to an optimal value as the response of the panel at resonance frequencies is damped down so that the low-frequency broadband sound transmission is monotonically reduced. References 40 to 42 present the development study of a smart panel with a 4×4 array of decentralized MIMO direct velocity control units which are formed by a square piezoelectric patch actuator with an accelerometer sensor at its center. With this arrangement quite good damping levels have been obtained in the low-frequency resonances of the panel, which have produced reductions of the averaged vibration of the panel between 5 and 8 dB for the third octave bands up to 1 kHz. However, the feedback control gains had to be limited to relatively low values since the decentralized control systems are only conditionally stable as the accelerometer sensor and piezoelectric actuator behaves as a collocated and dual pair only at low frequencies such that the bending wavelength is larger than the dimensions of the piezoelectric patch.⁴²

In this paper, we present a new type of smart panel with sixteen decentralized active damping control units which are composed of triangularly shaped piezoelectric patch actuators, with the base edges aligned along the borders of the panel, and accelerometer sensors placed at the vertices opposite to the base edges of the actuators. The sixteen control units are evenly spaced along the four edges of the panel so that they both look like and work like “active wedges.” The triangular shape of the actuator generates bending moments along the edges plus a transverse point force at the vertex opposite to the base edge⁴³ which results to be collocated and dual with the linear velocity measured by the accelerometer sensor. This makes the sensor actuator pair able to have better collocation and duality features than the one with the square piezoelectric patch actuator. As a result, larger control gains could be implemented in a direct velocity feedback loop which should then produce larger active damping.⁷ Thus, these control units can indeed be considered to be a sort of structural active wedges whose damping effect reduces reflections of incident flexural waves to the edges of the panel which could then be considered to be anechoic. In order to better highlight the control mechanisms and control effects of this type of smart panel with triangularly shaped piezoelectric actuators, the theoretical study presented in this paper contrasts its flexural response and sound radiation with those of a smart panel with a 4×4 array of MIMO-decentralized direct velocity feedback control units evenly distributed over the panel surface which are formed by a square piezoelectric patch actuator with an accelerometer sensor at its center. Also, the effectiveness of simpler control

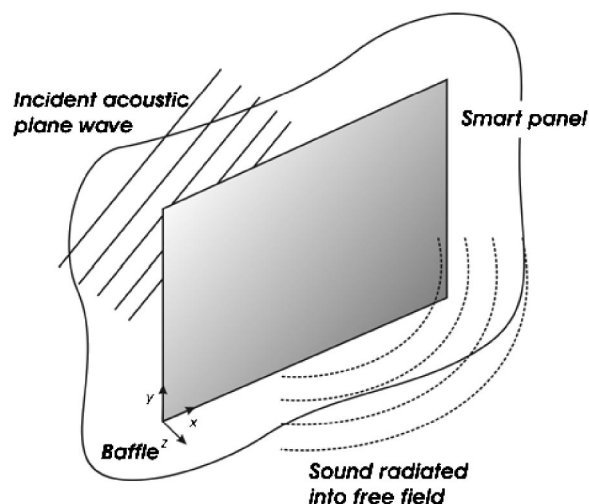


FIG. 1. Physical arrangement considered in a simulation study, in which the vibration of a simply supported panel is excited by a plane acoustic wave on one side and radiates sound into an anechoic half space on the other side of the panel.

configurations is considered for both types of smart panels where the sum of the sixteen velocity outputs is feedback to the sixteen actuators via a fixed gain SISO controller. The sensor-actuator response functions are studied both in frequency domain and with the Nyquist plots in order to assess the stability of the MIMO-decentralized and SISO control configurations.

The paper is structured into four parts. In Sec. II we present the analytical model used to predict the response and sound transmission through the smart panel when it is excited by a plane acoustic wave. In Sec. III the implementation of MIMO-decentralized or SISO direct velocity feedback control loops is modeled and discussed. Finally in Secs. IV and V the main control features of the two smart panels are analyzed and contrasted with reference to MIMO decentralized and SISO feedback control architectures, respectively.

II. RESPONSE AND SOUND TRANSMISSION THROUGH THE SMART PANELS

The steady state flexural response and sound radiation of smart panels which, as shown in Fig. 1, are excited by a harmonic acoustic plane wave, will be used as a model problem to illustrate the effects of the two sensor-actuator arrangements and control architectures shown in Fig. 2. The first arrangement is made by a 4×4 array of square piezoelectric patch actuators with at the centers velocity sensors evenly distributed over the surface of the panel. The second arrangement is instead made by sixteen triangularly shaped piezoelectric patch actuators, with the base edges evenly distributed along the perimeter of the panel, and velocity sensors at the vertices opposite to the base edges. The panel is made of aluminum, with dimensions $l_{xp} \times l_{yp} = 278 \times 247$ mm and thickness $h_p = 1$ mm and it is assumed to be baffled and simply supported along the perimeter. The material and geometrical properties of the panel are summarized in Table I. The far-field sound radiation is determined by the time-averaged total sound power radiated by the panel while

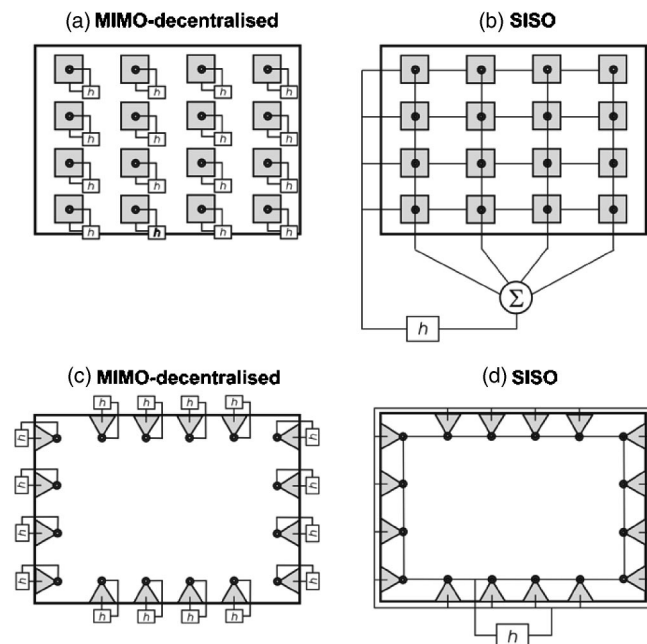


FIG. 2. Smart panels with a 4×4 grid of closely located ideal velocity sensors and square piezoelectric patch actuators or sixteen closely located ideal velocity sensors and triangular piezoelectric patch actuators with the base edges evenly aligned along the four edges of the panel which are arranged either in MIMO-decentralized (type a and c) or SISO (type b and d) direct feedback loops.

its time-averaged total kinetic energy is used to represent the near-field sound radiation which is determined by both strong and weakly radiating modes.¹⁶

The square piezoelectric patches, which have dimensions $a_x \times a_y = 25 \times 25$ mm and thickness $h_{\text{PZT}} = 0.2$ mm, are taken to be evenly distributed over the panel surface so that the distance between the centers of two adjacent patches is given by $d_{sx}, d_{sy} = l_{xp}/4, l_{yp}/4 = 69.5, 61.7$ mm and the distances of the centers of the external patches with the borders of the panel are given by $d_{sx}/2, d_{sy}/2 = 34.8, 30.9$ mm. The square piezoelectric patches, which have base and height dimensions $b, a = 40, 25$ mm and thickness $h_{\text{PZT}} = 0.2$ mm, are instead taken to be evenly distributed along the edges of the panel so that the distances between the centers of the base edges aligned along the $y = 0, l_{yp}$ borders of the panel are given by $d_{tx} = (l_{xp} - 2a)/5 = 45.6$ mm while the distances

TABLE I. Geometry and physical parameters for the panel.

Parameter	Value
Dimensions	$l_{xp} \times l_{yp} = 278 \times 247$ mm
Thickness	$h_p = 1$ mm
Mass density	$\rho_p = 2720$ Kg/m ³
Young's modulus	$E_p = 7 \times 10^{10}$ N/m ²
Poisson ratio	$\nu_p = 0.33$
Modal damping ratio	$\zeta_n = 0.02$

between the centers of the base edges aligned along the $x = 0, l_{xp}$ borders of the panel are given by $d_{ty} = (l_{yp} - 2a)/3 = 65.7$ mm. Also, the centers of the base edges of the outer patches aligned along the $y = 0, l_{yp}$ borders of the panel are located at a distance $d_{ty} + a = 70.6$ mm from the $x = 0, l_{xp}$ border while the centers of the base edges of the outer patches aligned along the $x = 0, l_{xp}$ borders of the panel are located at a distance $a = 25$ mm from the $y = 0, l_{yp}$ border. In this way, as shown in the two bottom schemes of Fig. 2, the vertices of the sixteen triangular patches define a rectangle of dimensions $(l_{xp} - 2a) \times (l_{yp} - 2a)$. The material and geometrical properties of the square and triangular piezoelectric, PZT (lead zirconate titanate), patches are summarized in Table II.

The steady state response of the panel has been derived assuming the incident acoustic plane wave to be harmonic with time dependence of the form $\text{Re}\{\exp(j\omega t)\}$ where ω is the circular frequency and $j = \sqrt{-1}$. The mechanical and electrical functions used in the model have therefore been taken to be the real part of counterclockwise rotating complex vectors, e.g., phasors, given in the form $v(\omega)e^{j\omega t}$ where $v(\omega)$ is the phasor at $t = 0$. A harmonic acoustic wave is assumed to excite the panel with azimuthal and elevation angles of $\phi = 45^\circ$ and $\theta = 45^\circ$ so that it excites all the structural modes of the panel. The sound pressure field $p_i(x, y, t)$ in the plane of the plate is therefore given by

$$p_i(x, y, t) = \text{Re}\{p_i(\omega)e^{j(\omega t - k_x x - k_y y)}\}, \quad (1)$$

where $p_i(\omega)$ is the phasor of the incident wave, $k_x = k_0 \sin(\theta)\cos(\phi)$ and $k_y = k_0 \sin(\theta)\sin(\phi)$ are the acoustic wave numbers in x - and y -directions where $k_0 = \omega/c_0$ and $c_0 = 343$ m/s are, respectively, the acoustic wave number and speed of sound in air.

TABLE II. Geometry and physical parameters for the piezoelectric, PZT (lead zirconate, titanate), patches.

Parameter	Value
Dimensions of the square patches	$a_x \times a_y = 25 \times 25$ mm
Base and height dimensions of the triangular patches	$b, a = 40, 25$ mm
Distances between the centers of two square patches	$d_{sx} = 69.5$ mm, $d_{sy} = 61.7$ mm
Distances between the centers of the base edges aligned either along the $y = 0, l_{yp}$ or $x = 0, l_{xp}$ borders of the panel	$d_{tx} = 45.6$ mm, $d_{ty} = 65.7$ mm
Thickness of the square and triangular patches	$h_{\text{PZT}} = 0.2$ mm
Density	$\rho_{\text{PZT}} = 7600$ kg/m ³
Young's modulus	$E_{\text{PZT}} = 6.3 \times 10^{10}$ N/m ²
Poisson ratio	$\nu_{\text{PZT}} = 0.29$
PZT strain/charge constants	$d_{31}^o = 166 \times 10^{-12}$ m/V $d_{32}^o = 166 \times 10^{-12}$ m/V $d_{36}^o = 0$

Both the time-averaged total kinetic energy and time-averaged total sound power radiated have been derived by dividing the panel into a grid of quadrilateral elements whose dimensions have been taken to be $l_{xe} = l_{xp}/(4N_1)$ and $l_{ye} = l_{yp}/(4N_2)$, where N_1 and N_2 are the higher modal orders used in the calculations. The phasors of the complex transverse velocities, $\dot{w}_{er}(\omega)$, at the centers of these R elements have been grouped into the following column vector:

$$\mathbf{w}_e(\omega) \equiv \begin{Bmatrix} \dot{w}_{e1}(\omega) \\ \dot{w}_{e2}(\omega) \\ \vdots \\ \dot{w}_{eR}(\omega) \end{Bmatrix}. \quad (2)$$

The flexural vibration of the smart panel under study is given by the superposition of the acoustic primary excitation generated by the incident plane wave and the structural secondary excitations generated by the control transducers bonded on the panel. Thus, assuming the system is linear, and assuming the radiated pressure has no effect on the panel vibration, then the vector with the phasors of the velocities at the centers of the elements, $\mathbf{w}_e(\omega)$, can be derived with the following matrix relation:

$$\mathbf{w}_e(\omega) = \mathbf{Y}_{ep}(\omega) p_i(\omega) + \mathbf{Y}_{ec}(\omega) \mathbf{v}_c(\omega), \quad (3)$$

where $\mathbf{v}_c(\omega)$ is a vector with the phasors of the complex input voltage signals, $v_{cs}(\omega)$, to the S control piezoelectric transducers:

$$\mathbf{v}_c(\omega) \equiv \begin{Bmatrix} v_{c1}(\omega) \\ v_{c2}(\omega) \\ \vdots \\ v_{cS}(\omega) \end{Bmatrix}. \quad (4)$$

The elements in the two matrices of Eq. (3) have been derived with a finite modal expansion so that⁴⁴

$$Y_{ep}^{r,1}(\omega) = j\omega \sum_{n=1}^N \frac{\psi_n(x_r, y_r) F_{np}(\omega)}{\rho_p h_p l_{xp} l_{yp} (\omega_n^2 - \omega^2 + j2\zeta_n \omega \omega_n)}, \quad (5)$$

$$Y_{ec}^{r,s}(\omega) = j\omega \sum_{n=1}^N \frac{\psi_n(x_r, y_r) F_{nc,s}(\omega)}{\rho_p h_p l_{xp} l_{yp} (\omega_n^2 - \omega^2 + j2\zeta_n \omega \omega_n)}, \quad (6)$$

where ρ_p and h_p are the density and the thickness of the smart panel ζ_n is the modal damping ratio, which was taken to be 0.01 for all modes in these simulations, ω_n and $\psi_n(x, y)$ are, respectively, the n -th natural frequency and natural mode, which for a simply supported panel are given by

$$\omega_n = \sqrt{\frac{D_p}{\rho_p h_p} \left[\left(\frac{n_1 \pi}{l_{xp}} \right)^2 + \left(\frac{n_2 \pi}{l_{yp}} \right)^2 \right]}, \quad (7)$$

$$\psi_n(x, y) = 2 \sin\left(\frac{n_1 \pi x}{l_{xp}}\right) \sin\left(\frac{n_2 \pi y}{l_{yp}}\right), \quad (8)$$

where $D_p = E_p h_p^3 / 12(1 - \nu_p^2)$ is the bending stiffness of the smart panel, with E_p and ν_p the Young's modulus of elasticity and Poisson's ratio, and n_1, n_2 are the two modal integers for the n -th mode. Finally $F_{np}(\omega)$ and $F_{nc,s}(\omega)$ are the two modal excitation terms which are due, respectively, to the primary acoustic excitation generated by the incident plane

wave and to the secondary flexural excitations generated by either the s -th square or s -th triangular piezoelectric patch control actuators bonded on the panel. The modal primary excitation terms, $F_{np}(\omega)$, are obtained by integrating the pressure field generated over the panel surface, that is⁴⁵

$$F_{np}(\omega) = p_i(\omega) \int_0^{l_{xp}} \int_0^{l_{yp}} \psi_n(x, y) e^{-j(k_x x + k_y y)} dx dy \\ = 4 p_i(\omega) I_{n1} I_{n2} l_{xp} l_{yp}, \quad (9)$$

where, if $n_1 \pi \neq \pm \sin \theta \cos \phi (\omega l_{xp} / c_0)$ and $n_2 \pi \neq \pm \sin \theta \sin \phi (\omega l_{yp} / c_0)$,

$$I_{n1} = \frac{n_1 \pi [1 - (-1)^{n_1} e^{-j \sin \theta \cos \phi (\omega l_{xp} / c_0)}]}{[n_1 \pi]^2 - [\sin \theta \cos \phi (\omega l_{xp} / c_0)]^2} \quad (10a)$$

and

$$I_{n2} = \frac{n_2 \pi [1 - (-1)^{n_2} e^{-j \sin \theta \sin \phi (\omega l_{yp} / c_0)}]}{[n_2 \pi]^2 - [\sin \theta \sin \phi (\omega l_{yp} / c_0)]^2}; \quad (10b)$$

and, if $n_1 \pi = \pm \sin \theta \cos \phi (\omega l_{xp} / c_0)$ and $n_2 \pi = \pm \sin \theta \sin \phi (\omega l_{yp} / c_0)$,

$$I_{n1} = (j/2) \operatorname{sgn}(\sin \theta \cos \phi) \quad (11a)$$

and

$$I_{n2} = (j/2) \operatorname{sgn}(\sin \theta \sin \phi). \quad (11b)$$

The modal secondary excitation terms, $F_{nc,s}(\omega)$, are derived by integrating the bending excitation fields generated by the control actuators over the panel surface. According to Ref. 46 and as shown in Fig. 3(a), if the principal axes x', y', z' of the piezoelectric material are aligned along the x, y, z axes of the panel, the s -th square piezoelectric patch actuator produces moment excitations,

$$m_{xs}(x_{es1,3}, y_{es1,3}, t) = \pm \frac{h_s}{2} e_{32}^0 v_{cs}(t), \quad (12a)$$

$$m_{ys}(x_{es2,4}, y_{es2,4}, t) = \pm \frac{h_s}{2} e_{31}^0 v_{cs}(t), \quad (12b)$$

respectively along the horizontal edges 1, 3, with coordinates $x_{cs} - a_x/2 \leq x_{es1,3} \leq x_{cs} + a_x/2$ and $y_{es1,3} = y_{cs} \mp a_y/2$, and along the vertical edges 2, 4, with coordinates $x_{es2,4} = x_{cs} \pm a_x/2$ and $y_{cs} - a_y/2 \leq y_{es2,4} \leq y_{cs} + a_y/2$. The indices 1 to 4 indicate the four edges in anticlockwise order starting from the bottom horizontal edge and, as given in Table II, a_x, a_y are the dimensions of the patches while x_{cs}, y_{cs} are the center coordinates of the s -th patch. Also, as shown in Fig. 3(a), point forces,

$$f_z(x_{vs2,4}, y_{vs2,4}, t) = -f_z(x_{vs1,3}, y_{vs1,3}, t) = \frac{h_s}{2} e_{36}^0 v_{cs}(t), \quad (12c)$$

are exerted at the four vertices of coordinates $(x_{vs1,3}, y_{vs1,3}) = (x_{cs} \mp a_x/2, y_{cs} \mp a_y/2)$, and $(x_{vs2,4}, y_{vs2,4}) = (x_{cs} \pm a_x/2, y_{cs} \mp a_y/2)$ of each patch. In Eqs. (12a)–(12c), h_s is the total thickness of the panel and piezoelectric patch, that is $h_s = h_p + h_{pZT}$. The piezoelectric stress/charge parameters, e_{31}^0, e_{32}^0 and e_{36}^0 , are derived from the following relation:⁴⁶

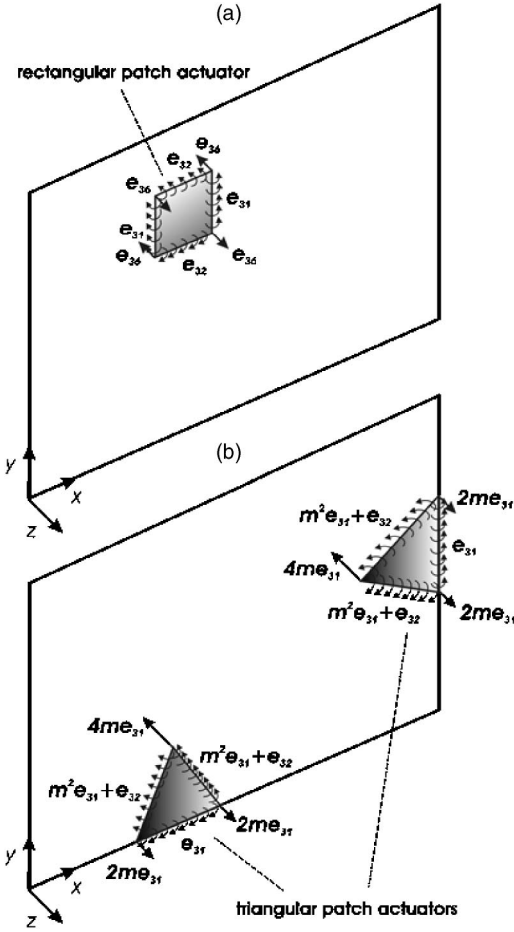


FIG. 3. Panels with square (top) or triangular (bottom) piezoelectric patches.

$$\begin{Bmatrix} e_{31}^0 \\ e_{32}^0 \\ e_{36}^0 \end{Bmatrix} = \begin{bmatrix} \frac{E_{PZT}}{(1-\nu_{PZT}^2)} & \frac{\nu_{PZT}E_{PZT}}{(1-\nu_{PZT}^2)} & 0 \\ \frac{\nu_{PZT}E_{PZT}}{(1-\nu_{PZT}^2)} & \frac{E_{PZT}}{(1-\nu_{PZT}^2)} & 0 \\ 0 & 0 & \frac{E_{PZT}}{2(1+\nu_{PZT})} \end{bmatrix} \times \begin{Bmatrix} d_{31}^0 \\ d_{32}^0 \\ d_{36}^0 \end{Bmatrix}, \quad (13)$$

where the Young's modulus of elasticity, E_{PZT} , the Poisson's ratio, ν_{PZT} , and the piezoelectric strain/charge, d_{31}^0 , d_{32}^0 and d_{36}^0 , parameters of the piezoelectric PZT (lead zirconate, titanate) material considered in this paper are given in Table II. Since the piezoelectric material considered in this paper has $e_{36}^0=0$, then there are no force excitations at the four vertices. Therefore the secondary modal excitation terms for square piezoelectric patch actuators, $F_{nc,s}(\omega)$, are given by the sum of the integrals along the four edges of the piezoelectric patch of the first derivatives of the panel natural modes in directions orthogonal to the edges, pointing outside the square surface, multiplied by the appropriate excitation coefficients given in Eqs. (12a), (12b) that is,

$$\begin{aligned} F_{nc,s}(\omega) = & \frac{h_s}{2} e_{32}^0 \left\{ + \int_{x_{vs1}}^{x_{vs2}} \frac{\partial \psi_n(x, y_{es1})}{\partial y} dx \right. \\ & - \left. \int_{x_{vs3}}^{x_{vs4}} \frac{\partial \psi_n(x, y_{vs3})}{\partial y} dx \right\} \\ & + \frac{h_s}{2} e_{31}^0 \left\{ \int_{y_{vs1}}^{y_{vs4}} \frac{\partial \psi_n(x_{es4}, y)}{\partial x} dy \right. \\ & - \left. \int_{y_{vs2}}^{y_{vs3}} \frac{\partial \psi_n(x_{es2}, y)}{\partial x} dy \right\}. \quad (14) \end{aligned}$$

The bending excitation field generated by a triangular piezoelectric patch has been derived in Ref. 43. Considering the triangular patches shown in Fig. 3(b), which are bonded in such a way as their base edges, a , are aligned with either the x - or y -borders of the panel and assuming that the principal axes x', y', z' of the piezoelectric material are aligned in such a way as x' and y' are parallel, respectively, to the height a and base b of the triangular patches, then the s -th triangular piezoelectric patch actuator produce moment excitations along the base edge of amplitude,

$$m_{bs}(x_{bs}, y_{bs}, t) = \frac{h_s}{2} e_{31}^0 v_{cs}(t), \quad (15a)$$

and moment excitations along the two lateral edges of amplitude,

$$m_{ls1,2}(x_{ls1,2}, y_{ls1,2}, t) = \frac{h_s}{2} (m^2 e_{31}^0 + e_{32}^0) v_{cs}(t). \quad (15b)$$

The positions of the triangular patches have been defined with reference to the middle points of their base edges x_{ms} and y_{ms} . Thus the coordinates of the base edge of the s -th triangular piezoelectric patch actuator are either $x_{ms} - b/2 \leq x_{bs} \leq x_{ms} + b/2$, $y_{bs} = 0, l_{yp}$ or $x_{bs} = 0, l_{xp}$, $y_{ms} - b/2 \leq y_{bs} \leq y_{ms} + b/2$ depending whether the base is aligned along the $y = 0, l_{yp}$ or $x = 0, l_{xp}$ borders of the panel. Also, the coordinates of the lateral edges of the s -th triangular piezoelectric patch actuator with the base aligned along the $y = 0, l_{yp}$ borders of the panel are, respectively, $y_{ls1,2} = \pm m[x_{ls1,2} - (x_{ms} \mp b/2)]$ and $y_{ls1,2} = \mp m[x_{ls1,2} - (x_{ms} \mp b/2)]$ with $x_{ms} - b/2 \leq x_{ls1} \leq x_{ms}$ and $x_{ms} \leq x_{ls2} \leq x_{ms} + b/2$. Finally the coordinates of the lateral edges of the s -th triangular piezoelectric patch actuator with the base aligned along the $x = 0, l_{xp}$ borders of the panel are, respectively, $x_{ls1,2} = \pm m[y_{ls1,2} - (y_{ms} \mp b/2)]$ and $x_{ls1,2} = \mp m[y_{ls1,2} - (y_{ms} \mp b/2)]$ with $y_{ms} - b/2 \leq y_{ls1} \leq y_{ms}$ and $y_{ms} \leq y_{ls2} \leq y_{ms} + b/2$. As listed in Table II, b, a are, respectively, the base and high of the triangular patch and $m = b/2a$ is the slope of the lateral edges. Finally as found for the square piezoelectric patches $h_s = h_p + h_{PZT}$. Also three point forces are generated at the vertices of the s -th triangular piezoelectric patch actuator,

$$f_z(x_{vs1,2}, y_{vs1,2}, t) = 2m \frac{h_s}{2} e_{31}^0 v_{cs}(t), \quad (15c)$$

$$f_z(x_{vs3}, y_{vs3}, t) = -4m \frac{h_s}{2} e_{31}^0 v_{cs}(t), \quad (15d)$$

where the vertices of the s -th triangular piezoelectric patch actuators with the base aligned along the $y=0, l_{yp}$ borders of the panel are, respectively, $x_{vs1,2}=x_{ms} \mp b/2$, $y_{vs1,2}=0, l_y$ and $x_{vs3}=x_{ms}$, $y_{vs3}=a, (l_y - a)$ while the vertices of the s -th triangular piezoelectric patch actuators with the base aligned along the $x=0, l_{xp}$ borders of the panel are, respectively, $x_{vs1,2}=0, l_x$, $y_{vs1,2}=y_{ms} \mp b/2$ and $x_{vs3}=a, (l_x - a)$, $y_{vs3}=y_{ms}$. In summary the secondary modal excitation terms for triangularly shaped piezoelectric patch actuators with the base edge aligned along the perimeter of the panel, $F_{nc,s}(\omega)$, are given by the sum of the integrals along the three edges of the piezoelectric patch of the first derivatives of the natural modes in directions orthogonal to the edges, pointing outside the triangular surface, plus the amplitudes of the modes at the tip vertices multiplied by the appropriate excitation coefficients given, respectively, in Eqs. (15a), (15b), (15d), that is,

$$F_{nc,s}(\omega) = \frac{h_s}{2} (m^2 e_{31}^0 + e_{32}^0) \left\{ \int_{v_{s1}}^{v_{s3}} \frac{\partial \psi_n(x,y)}{\partial \mathbf{n}_{ls1}} dl_{s1} + \int_{v_{s2}}^{v_{s3}} \frac{\partial \psi_n(x,y)}{\partial \mathbf{n}_{ls2}} dl_{s2} \right\} + \frac{h_s}{2} e_{31}^0 \int_{v_{s1}}^{v_{s2}} \frac{\partial \psi_n(x,y)}{\partial \mathbf{n}_{bs}} db_s - 4 \frac{h_s}{2} m e_{31}^0 \psi_n(x_{sv3}, y_{sv3}), \quad (16)$$

where $v_{sj}=(x_{vsj}, y_{vsj})$ indicates the coordinates of the vertices of the triangular patch as given above and $\mathbf{n}_{ls1}, \mathbf{n}_{ls2}, \mathbf{n}_{bs}$ are the normal unit vectors to the lateral and base edges pointing outside the triangular surface of the actuator. Equation (16) does not account for the two forces acting at the base vertices of the triangular patches because the panel is simply supported and therefore does not allow transverse excitations along its perimeter.

The time-averaged total kinetic energy of the panel is given by

$$E(\omega) = \frac{\rho_p h_p}{4} \int_0^{l_{xp}} \int_0^{l_{yp}} |\dot{w}(x,y,\omega)|^2 dx dy, \quad (17)$$

where $\dot{w}(x,y,\omega)$ is the phasor of the transverse velocity over the panel surface. This expression can be approximated by the summation of the kinetic energies of each element into which the panel has been subdivided so that

$$E(\omega) = \frac{M_e}{4} \mathbf{w}_e^H(\omega) \mathbf{w}_e(\omega), \quad (18)$$

where $M_e = \rho_p h_p l_{xe} l_{ye}$ is the mass of each element and H denotes the Hermitian transpose. The time-averaged total sound power radiation by a baffled panel can be derived by integrating the product of the phasor of the nearfield sound pressure, $p_0(x,y,\omega)$, on the radiating surface and the phasor of the transverse velocity of the panel, $\dot{w}(x,y,\omega)$, so that

$$W_r(\omega) = \frac{1}{2} \int_0^{l_{xp}} \int_0^{l_{yp}} \text{Re}[\dot{w}(x,y,\omega)^* p_0(x,y,\omega)] dx dy, \quad (19)$$

where $*$ denotes the complex conjugate. For the baffled flat plate considered in this paper, the acoustic pressure $p_0(x,y,\omega)$ can be written in terms of the surface velocity using the Rayleigh integral¹

$$p_0(x,y,\omega) = \frac{j\omega\rho_0}{2\pi} \int_0^{l_{xp}} \int_0^{l_{yp}} \dot{w}(x',y',\omega) \frac{e^{-jk_0 r}}{r} dx' dy', \quad (20)$$

where $r = \sqrt{(x-x')^2 + (y-y')^2}$ is the distance between the point (x,y) where the sound pressure is estimated and vibration velocity positions on the panel (x',y') and $\rho_0 = 1.21 \text{ kg/m}^3$ is the density of air. Substituting Eq. (20) in (19), the time average total sound radiation is found to be given by a quadruple integral:

$$W_r(\omega) = \frac{\omega\rho_0}{4\pi} \int_0^{l_{xp}} \int_0^{l_{yp}} \int_0^{l_{xp}} \int_0^{l_{yp}} \dot{w}(x,y,\omega)^* \dot{w}(x',y',\omega) \times \frac{\sin k_0 r}{r} dx' dy' dx dy. \quad (21)$$

The quadruple integral in Eq. (21) can also be approximated by summing the radiation contributions of all the elements into which the panel has been subdivided, so that the time-averaged total sound power radiation can be expressed as¹⁶

$$W_r(\omega) = \frac{A_e}{2} \text{Re}[\mathbf{w}_e^H(\omega) \mathbf{p}_e(\omega)], \quad (22)$$

where $A_e = l_{xe} l_{ye}$ is the area of each element and $\mathbf{p}_e(\omega)$ is the vector with the phasors of the sound pressure terms in front of the panel at the center positions of the grid of elements:

$$\mathbf{p}_e(\omega) \equiv \begin{Bmatrix} p_{e1}(\omega) \\ p_{e2}(\omega) \\ \vdots \\ p_{eR}(\omega) \end{Bmatrix}. \quad (23)$$

Following Ref. 16, Eq. (12) can also be written as

$$W_r(\omega) = \frac{A_e}{2} \text{Re}[\mathbf{w}_e^H(\omega) \mathbf{Z}(\omega) \mathbf{w}_e(\omega)] = \mathbf{w}_e^H(\omega) \mathbf{R}(\omega) \mathbf{w}_e(\omega), \quad (24)$$

where $\mathbf{Z}(\omega)$ is the matrix with the point and transfer acoustic impedance terms over the grid of points into which the panel has subdivided:¹² $Z_{ij}(\omega) = (j\omega\rho_0 A_e / 2\pi r_{ij}) e^{-jk_0 r_{ij}}$, with r_{ij} the distance between the centers of the i -th and j -th elements. The matrix \mathbf{R} is defined as the radiation matrix which is given by¹²

$$\mathbf{R}(\omega) = \frac{A_e}{2} \text{Re}[\mathbf{Z}(\omega)] = \frac{\omega^2 \rho_0 A_e^2}{4 \pi c_0} \begin{bmatrix} 1 & \frac{\sin(k_0 r_{12})}{k_0 r_{12}} & \dots & \frac{\sin(k_0 r_{1R})}{k_0 r_{1R}} \\ \frac{\sin(k_0 r_{21})}{k_0 r_{21}} & 1 & & \\ \dots & \dots & \dots & \dots \\ \frac{\sin(k_0 r_{R1})}{k_0 r_{R1}} & & & 1 \end{bmatrix}. \quad (25)$$

Since the primary excitation is an acoustic wave, the ratio of the time-averaged total sound power radiated, $W_r(\omega)$, to the time-averaged incident sound power, $W_i(\omega)$, which is termed the sound transmission ratio;

$$T(\omega) = W_r(\omega)/W_i(\omega), \quad (26)$$

has been used to describe the sound transmission phenomenon. The time-averaged incident sound power due to the plane acoustic wave is given by⁴⁵

$$W_i(\omega) = |p_i(\omega)|^2 l_{xp} l_{yp} \cos(\theta) / 2 \rho_0 c_0. \quad (27)$$

III. DIRECT VELOCITY FEEDBACK CONTROL

The phasors of the output error signal(s), $i_{ce}(\omega)$, from the E velocity error sensors can also be derived with a matrix relation of the type

$$\mathbf{i}_c(\omega) \equiv \mathbf{Y}_{cp}(\omega) p_i(\omega) + \mathbf{Y}_{cc}(\omega) \mathbf{v}_c(\omega), \quad (28)$$

where $\mathbf{i}_c(\omega)$ is the column vector with the phasors of the error sensor signal(s):

$$\mathbf{i}_c(\omega) \equiv \begin{Bmatrix} i_{c1}(\omega) \\ i_{c2}(\omega) \\ \vdots \\ i_{cE}(\omega) \end{Bmatrix}, \quad (29)$$

and $\mathbf{v}_c(\omega)$ is defined in (4). Assuming the velocity sensors used in the smart panels of Fig. 2 to be ideal transducers that measure the transverse velocity at the centers of the square patches or at the tips of the triangular patches, then the elements of the two matrices in Eq. (28) could be derived with a finite modal expansion considering the modal amplitude at the detection points $(x_e, y_e) = (x_{cs}, y_{cs})$ or, $(x_e, y_e) = (x_{vs3}, y_{vs3})$, respectively, for the square and triangular actuators, so that⁴⁴

$$Y_{cp}^r(\omega) = j\omega \sum_{n=1}^N \frac{\psi_n(x_s, y_s) F_{np}(\omega)}{\rho_p h_p l_{xp} l_{yp} (\omega_n^2 - \omega^2 + j2\zeta_n \omega \omega_n)}, \quad (30)$$

$$Y_{cc}^{r,s}(\omega) = j\omega \sum_{n=1}^N \frac{\psi_n(x_s, y_s) F_{nc,s}(\omega)}{\rho_p h_p l_{xp} l_{yp} (\omega_n^2 - \omega^2 + j2\zeta_n \omega \omega_n)}. \quad (31)$$

For the two panels types b and d in Fig. 2, where in order to implement a SISO velocity feedback loop the sensors outputs are summed up and the same control signal is feed to the control actuators, the total current output is still derived with Eq. (28) where the two mobility matrices \mathbf{Y}_{cp} and \mathbf{Y}_{cc} are pre-multiplied by a $1 \times E$ vector \mathbf{e} of unit terms and the matrix \mathbf{Y}_{cc} is also post-multiplied by an $S \times 1$ vector \mathbf{s} of unit

terms so that they become two scalar terms Y_{cp} and Y_{cc} .

The general block diagram of a multi-channel velocity feedback control system is shown in Fig. 4. If an equal number of actuators and velocity sensors is used, the matrix of plant responses, $\mathbf{Y}_{cc}(\omega)$, is square and the matrix of feedback control filters, $\mathbf{H}(j\omega)$, is also square. Provided the control system is stable, the vector with the phasors of the sensor(s) current output(s), $\mathbf{i}_c(\omega)$, is related to the phasor of the incident plane acoustic wave, $p_i(\omega)$, by the expression

$$\mathbf{i}_c(\omega) = [\mathbf{I} + \mathbf{Y}_{cc}(\omega) \mathbf{H}(\omega)]^{-1} \mathbf{Y}_{cp}(\omega) p_i(\omega). \quad (32)$$

Also the vector of control inputs to the actuators, $\mathbf{v}_c(\omega)$, is given by

$$\mathbf{v}_c(\omega) = -\mathbf{H}(\omega) [\mathbf{I} + \mathbf{Y}_{cc}(\omega) \mathbf{H}(\omega)]^{-1} \mathbf{Y}_{cp}(\omega) p_i(j\omega). \quad (33)$$

For single input single output control (SISO) the vectors and matrices reduce to scalars and thus the stability of the feedback control loop could be assessed using the classic feedback control theory.^{3,7,47} In particular, Balas⁹ has shown that, if the sensor-actuator pair is collocated and dual,⁸ then the SISO direct velocity feedback control loop is unconditionally stable. Indeed in this case the sensor-actuator frequency response function is real positive definite⁷ so that its Nyquist plot occupies the right hand side quadrants as ω varies from $-\infty$ to $+\infty$ and thus the Nyquist instability point $(-1 j0)$ is never encircled whatever is the control gain.

For multi-input multi-output (MIMO) decentralized control, $\mathbf{Y}_{cc}(\omega)$ is a fully populated matrix of input and transfer responses between the actuators and sensors on the panel and $\mathbf{H}(\omega)$ is a diagonal matrix which, for direct velocity feedback control, is assumed to have equal fixed gains so that $\mathbf{H}(\omega) = h \mathbf{I}$, where h is the feedback gain. In this case the stability

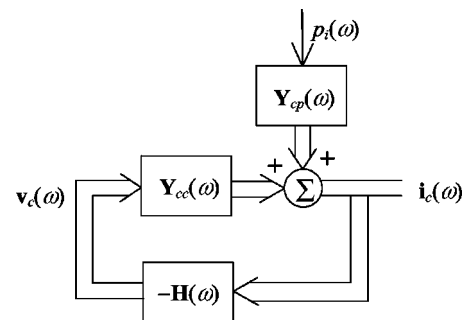


FIG. 4. Multichannel feedback control system, which for a passive plant response, $\mathbf{Y}_{cc}(j\omega)$, and a passive controller $\mathbf{H}(j\omega)$, is unconditionally stable.

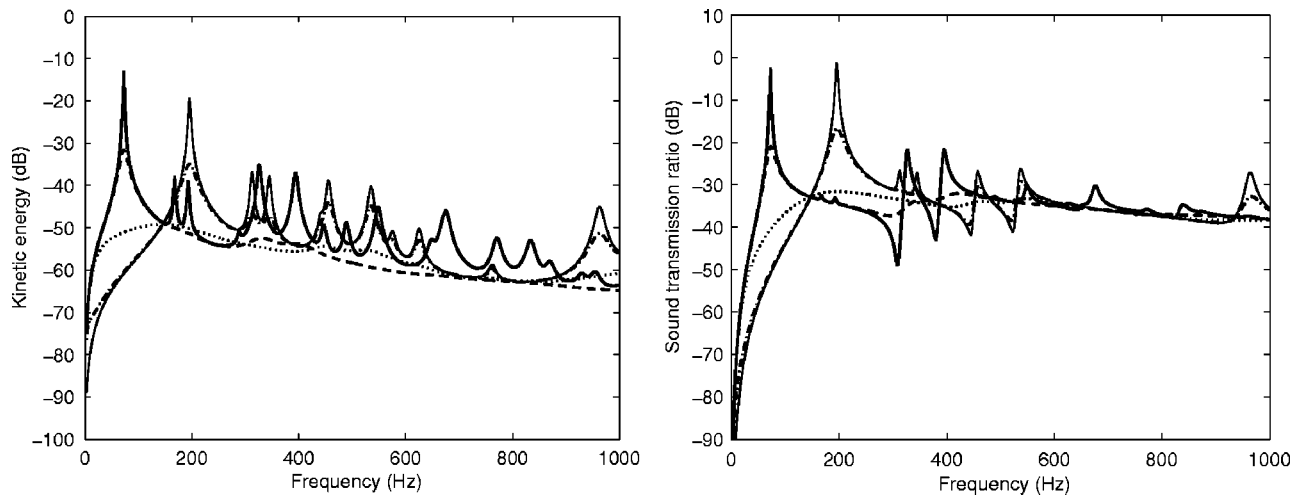


FIG. 5. Total kinetic energy (left) and sound transmission ratio (right) of the panel excited by a plane acoustic wave with no control (solid line) and with a 4×4 grid of closely located velocity sensors and square piezoelectric patches actuators (type a in Fig. 3) MIMO decentralized feedback controllers with feedback gains of 10 (dashed line), 10^2 (dotted line), 10^3 (dash-dotted line) and 10^4 (faint line).

of the MIMO decentralized control system can be determined by examining whether the locus of the determinant of $[\mathbf{I} + \mathbf{Y}_{cc}(j\omega)\mathbf{H}(j\omega)]$ encloses the origin^{48,49} as ω varies from $-\infty$ to $+\infty$. Alternatively the fact that the determinant of a matrix is the product of its eigenvalues can be used to derive a series of polar plots, each of which are analogous to the single channel Nyquist criteria. As found for the SISO control case, if collocated and compatible transducers are used,⁸ then the real part of $\mathbf{Y}_{cc}(\omega)$ must be positive definite and $\mathbf{Y}_{cc}(\omega)$ can be described as being passive. If $\mathbf{H}(\omega)$ is also passive, e.g., when it is equal to $h\mathbf{I}$ and $h > 0$, then the control system is unconditionally stable.^{39,48,49}

When the SISO or MIMO decentralized velocity feedback control systems are implemented, the total kinetic energy and sound transmission ratio given in Eqs. (18) and (24) can be derived after combining Eq. (3) with Eq. (33) so that

$$\mathbf{w}_e(\omega) = \mathbf{Y}_{ep}(\omega)p_i(\omega) - \mathbf{Y}_{ec}(\omega)\mathbf{H}(\omega)[\mathbf{I} + \mathbf{Y}_{cc}(\omega)\mathbf{H}(\omega)]^{-1} \times \mathbf{Y}_{cp}(\omega)p_i(j\omega). \quad (34)$$

It is important to underline that in the following sections the stability properties of the MIMO and SISO control systems in Fig. 2 will be discussed only at a qualitative level. Therefore the control effectiveness of the studied control systems are derived without taking into account whether it would be possible or not to implement the necessary gains without generating instabilities in the control loops.

IV. DECENTRALIZED MIMO DIRECT VELOCITY FEEDBACK CONTROL

In this section the control effectiveness of the smart panels with sixteen decentralized MIMO control units type (a) and type (c) in Fig. 2, which have either sixteen square piezoelectric patch actuators or sixteen triangular piezoelectric patch actuators, is investigated. The stability of the two decentralized control systems is also analyzed to some extent by considering Bode and Nyquist plots of the sensor-actuator frequency response function of one of the sixteen decentralized control systems in the two smart panels. Although the

stability of MIMO control systems should be assessed with reference to the locus of the determinant of $[\mathbf{I} + \mathbf{Y}_{cc}(j\omega)\mathbf{H}(j\omega)]$, for decentralized control, the stability of each control unit could also be evaluated independently using the classic feedback control theory^{3,7,47} which provides an indication of whether the decentralized MIMO control system is to be only conditionally stable. Moreover, if the smart panel is sufficiently damped and the control units are well separated from each others, then an indication about the gain or phase margins for each individual controller could also be derived.

A. Control effectiveness

The two plots in Fig. 5 show, respectively, the total kinetic energy and the sound transmission ratio of the smart panel with the 4×4 array of square piezoelectric patch actuators with the center velocity sensors' MIMO decentralized control system. The solid line on the left hand side plot for the total kinetic energy highlights the typical response of a panel which is characterized by a series of resonances whose amplitudes gradually roll off as the frequency rises. In particular the peaks of the first few resonances are relatively high and sharp because of the low damping effects at lower frequencies. The solid line on the right hand side plot for the sound transmission ratio shows a similar behavior, although there are almost no peaks for the resonances due to the even-even or even-odd natural modes of the panel which have relatively low sound radiation efficiency.¹

The dashed and dotted lines in the two plots of Fig. 5 show that as the gains of the sixteen control systems are raised the resonance peaks are flattened down. This is due to the active damping effect⁵⁰ generated by the sixteen DVFB control systems that indeed increase the overall damping of the lower-frequency resonant modes of the smart panel.^{39,40} However when relatively higher control gains are implemented this trend is inverted and, as shown by the dash-dotted and faint lines in the two plots of Fig. 5, the total kinetic energy and the sound transmission ratio are once more characterized by a new set of low-frequency reso-

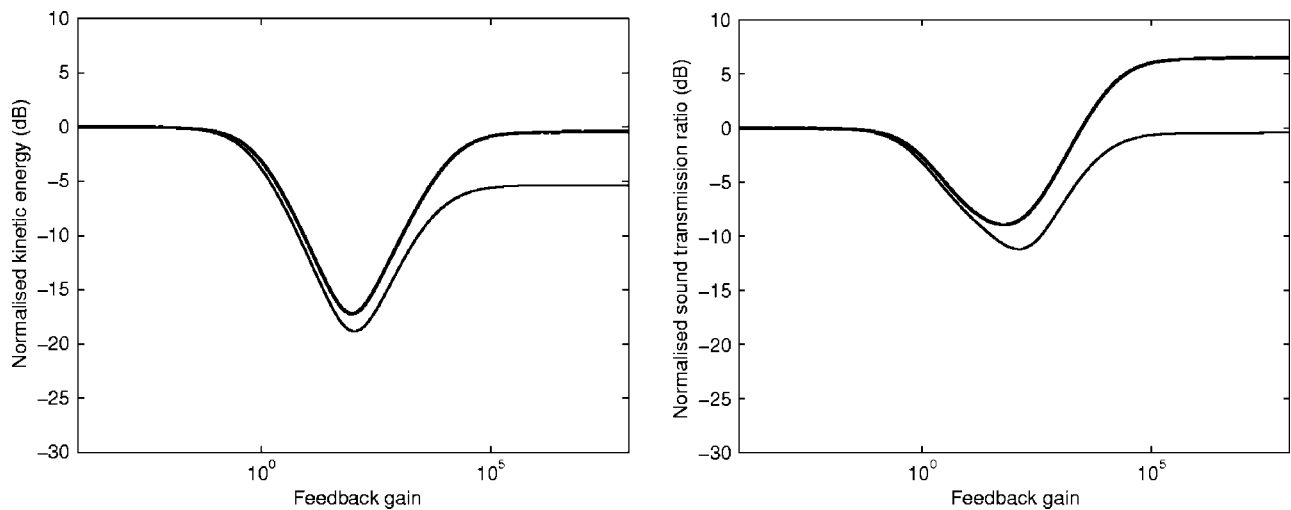


FIG. 6. Normalized total kinetic energy (left) and sound transmission ratio (right), integrated between 0 Hz and 1 kHz, plotted against the gain in the MIMO decentralized velocity feedback controllers, h , for the 4×4 grid of closely located velocity sensors and square piezoelectric patches actuators (solid line, type a in Fig. 2) and the 16 closely located velocity sensors and triangular piezoelectric patches actuators (faint line, type c in Fig. 2).

nances whose amplitudes are similar, if not higher, than those of the panel with no control. This is a typical control spillover phenomenon where, as discussed in Ref. 51, the large control gains produces a pinning effect at the control positions so that the response of the smart panel is rearranged into that of a lightly damped panel which is pinned at the sixteen control positions.

This type of behavior is summarized by the solid lines in the two plots of Fig. 6 which show how the normalized (normalized to the total kinetic energy when there is no control) total kinetic energy and normalized (normalized to the sound transmission ratio when there is no control) sound transmission ratio, integrated from 0 Hz to 1 kHz, vary with the feedback gain. Indeed both plots indicate that as the control gains are raised from zero as the frequency-averaged response and sound radiation of the smart panel monotonically falls down and reductions of the normalized total kinetic energy and normalized sound transmission ratio, respectively, of 17 dB and 9 dB could be achieved. If the control

gains are pushed farther up then, because of the pinning effect described above, the response of the smart panel is brought back to the levels with no control while its sound radiation is even increased by about 6 dB than in the case of no control. This is due to the fact that the new resonant modes of the smart panel generated by the pinning effects at the sixteen control positions have relatively higher sound radiation efficiency than the lower order modes of the unconstrained simply supported panel.¹

The two plots in Fig. 7 show the total kinetic energy and the sound transmission ratio of the smart panel with sixteen triangularly shaped piezoelectric patch actuators with velocity sensors at the vertices opposite to the base edges MIMO decentralized control system. From a qualitative point of view these two plots indicate that the sixteen control units with triangular actuators arranged along the perimeter of the panel produces similar effects than the sixteen control systems with square actuators distributed over the surface of the panel. The most important difference is found when rela-

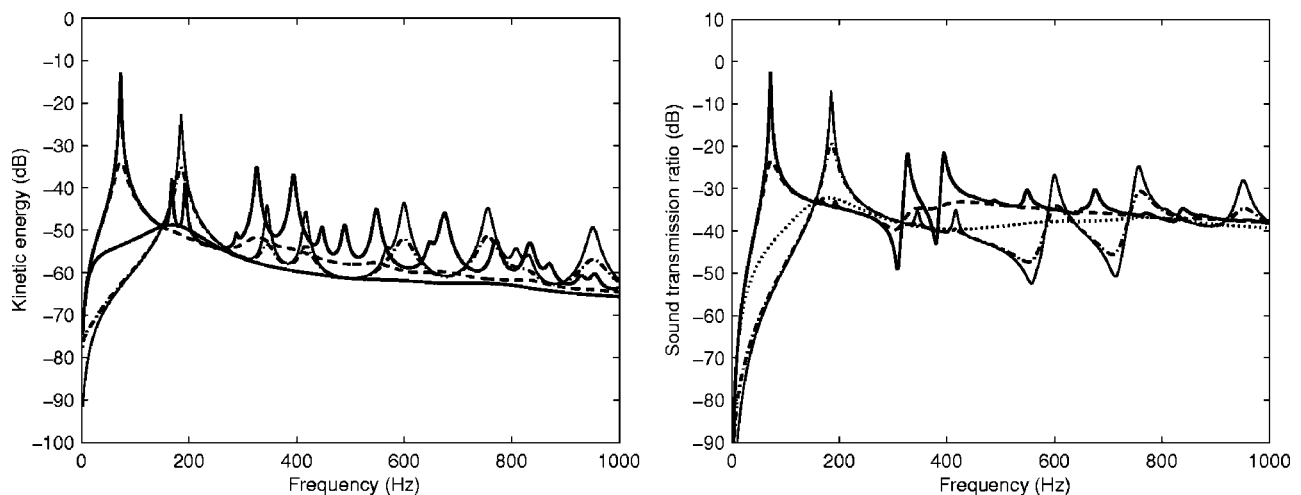


FIG. 7. Total kinetic energy (left) and sound transmission ratio (right) of the panel excited by a plane acoustic wave with no control (solid line) and with 16 closely located velocity sensors and 16 triangular piezoelectric patches actuators (type c in Fig. 2) MIMO decentralized feedback controllers with feedback gains of 10 (dashed line), 10^2 (dotted line), 10^3 (dash-dotted line) and 10^4 (faint line).

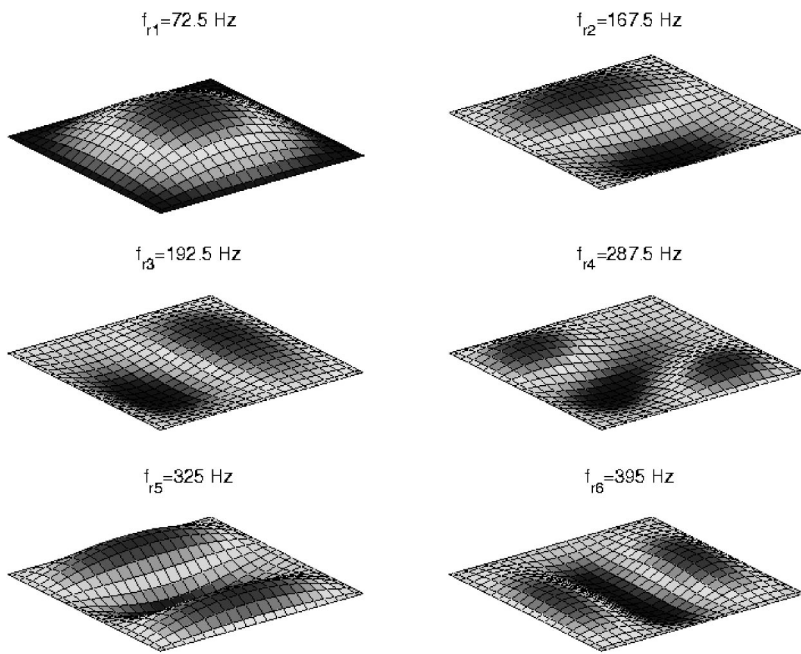


FIG. 8. Deflection shapes of the panel at the first six resonance frequencies with no control.

tively high control gains are implemented in which case the resonances of the new modes of the smart panel, which are generated by the pinning effect of the triangular actuators scattered along the perimeter of the panel, occur at relatively higher frequencies than those due to the pinning effect of the square actuators distributed over the surface of the panel.

Moving to a quantitative analysis the faint lines of the two plots in Fig. 6 indicates that the sixteen control systems with triangular actuators evenly distributed along the perimeter of the smart panel generates slightly larger control effects than those due to the square actuators evenly distributed over the surface of the smart panel. In fact the maximum reduction of the normalized total kinetic energy is increased from 17 dB to 19 dB while the maximum reduction of the normalized sound transmission ratio is increased from 9 dB to 11 dB. As for the smart panel with square actuators evenly distributed over the surface of the smart panel, when relatively large control gains are implemented, the pinning effects at the vertices of the triangular actuators generate a new set of lightly damped resonant modes, which, however, still produce a reduction of about 5 dB of the normalized kinetic energy since in this case the vibrating surface of the panel is reduced to virtual edges defined by the control positions at the tips of the triangular actuators arranged along the edges of the panel. The frequency-averaged sound radiation of the panel is instead brought back to the case with no control, probably because the reduction of the response of the smart panel is balanced by the increased radiation efficiencies of the new resonant modes generated by the pinning effects along the perimeter of the smart panel.

In order to analyze in detail the different behaviors of the two smart panels, the deflections shapes of the panels in correspondence to the first six resonances, which as shown in Fig. 8 are closely linked to the first four natural modes of the panel, have been considered. Figures 9 and 10 show how these six deflection shapes varies when either the optimal control gains, that give the best control effects, or when very

large control gains are implemented in the smart panels with either the sixteen decentralized control systems made by square actuators (left hand side plots) or the sixteen decentralized control systems made by triangular actuators (right hand side plots). Since the deflection shapes occur over a wide range of amplitudes, the plots in Figs. 8 to 10 have been normalized to have the same maximum deflections.

Comparing the two plots in Fig. 9 with that of Fig. 8, it is found that, when the optimal control gains are implemented, then the response of the panel at the first six resonance frequencies is generally modified in such a way that the deflection shapes of the smart panel are not anymore controlled by the co-respective natural modes of the panel. This confirms the active damping action which indeed tends to reduce the contribution of the resonant modes so that the residual response is controlled by nonresonant modes of the panel. It is important to note that the sixteen control units arranged along the perimeter of the smart panel produces the damping action exactly along the borders of the panel which, at frequencies below the critical frequency, are indeed the portions of the panel which generates the sound radiation.¹ Thus the idea of scattering the control units along the perimeter of the smart panel is not convenient just for control stability issues, as discussed in the following section, or for practical matters such as the fact that the central part of the panel is not occupied by the control systems which are instead located near the borders of the panel where it is much easier to arrange the electronics of the sixteen control systems. On the contrary the control units located along the perimeter of the panel generate active damping exactly over the portion of surface of the smart panel that primarily causes sound radiation. Thus the triangularly shaped control actuators arranged along the edges of the panel could indeed be referred as “active structural wedges” that reduce the reflection of incident flexural waves. It is therefore realistic to presume that when localized structural excitations are generated on the panel then even bigger control effects should be

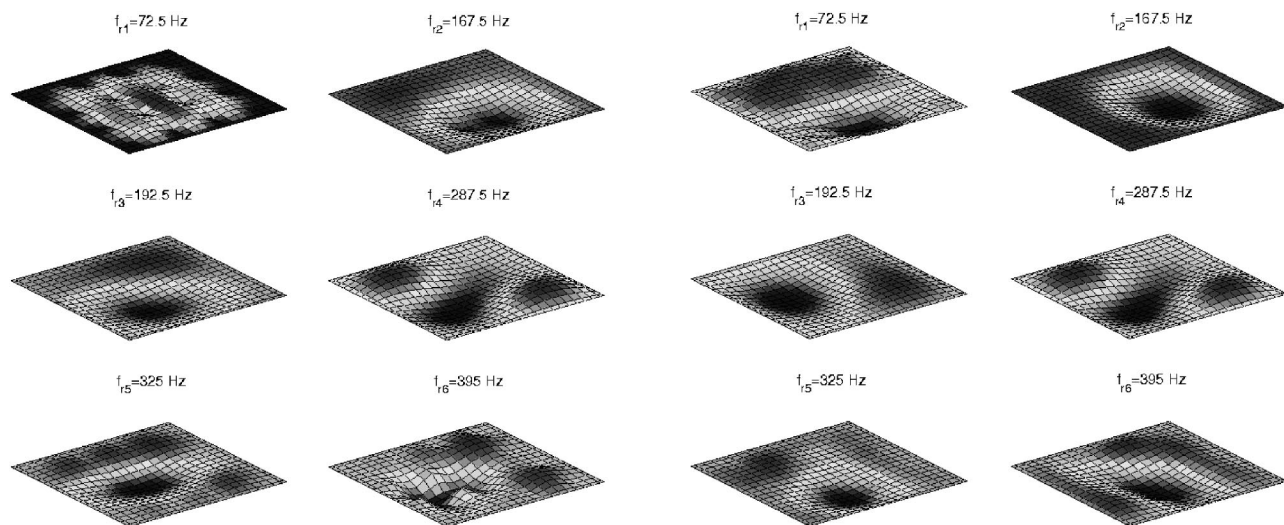


FIG. 9. Deflections shapes of the panel at the first six resonance frequencies when the optimal feedback control gains are implemented in the 4×4 grid of closely located velocity sensors and square piezoelectric patches actuators (left, type a in Fig. 2) and the 16 closely located velocity sensors and triangular piezoelectric patches actuators (right, type c in Fig. 2).

obtained than those found for the incident acoustic wave disturbance.

When very large control gains are implemented in the sixteen control loops then, as shown in the two plots of Fig. 10, the deflection shapes relative to the first six new resonance frequencies clearly show the pinning actions of the control units which, as shown on the left hand side plot, occurs on the 4×4 grid of control points for the system with the square piezoelectric patch actuators or, as shown on the right hand side plot, are located along the perimeter of the panel for the system with the triangular piezoelectric patch actuators. As highlighted by Fahy,¹ a periodically supported panel better radiates sound than a one bay panel since the periodic constraints generates new “edges” around which extra sound is radiated. This is why the frequency-averaged sound radiation of the smart panel constrained by the 4×4 grid of control systems with the square piezoelectric patch

actuators that implement large control gains is about 5 dB higher than that of the unconstrained one-bay panel. In contrast, the sixteen control units with triangular actuators that implement large control gains produce a pinning effect around the edges of the panel such that the deflection shapes are still characterized by a central part which is not constrained. As a result the frequency-averaged sound radiation is about the same to that of the one-bay panel. Essentially, the right hand side plot in Fig. 6, indicates that it is just a little lower probably because the radiating surface of the actively constrained smart panel is smaller than that of the unconstrained panel.

B. Control stability of a single control unit

The stability properties of the two independent control units in Figs. 2(a) and 2(c) are examined using the classic

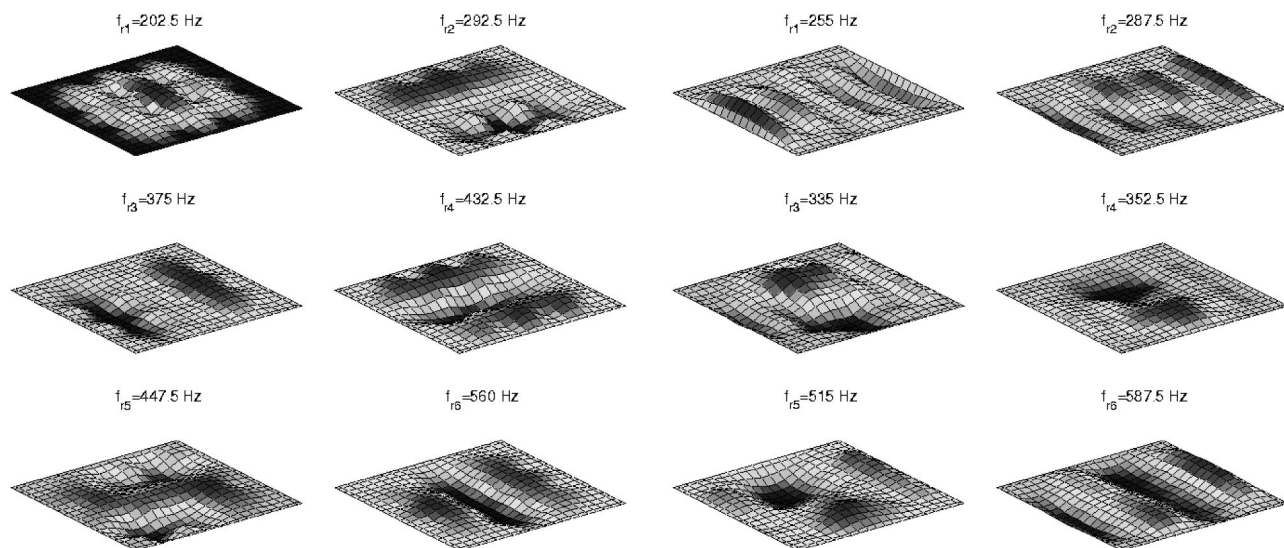


FIG. 10. Deflection shapes of the panel at the first six new resonance frequencies generated by very large feedback gains in the 4×4 grid of closely located velocity sensors and square piezoelectric patches actuators (left, type a in Fig. 2) and the 16 closely located velocity sensors and triangular piezoelectric patches actuators (right, type c in Fig. 2).

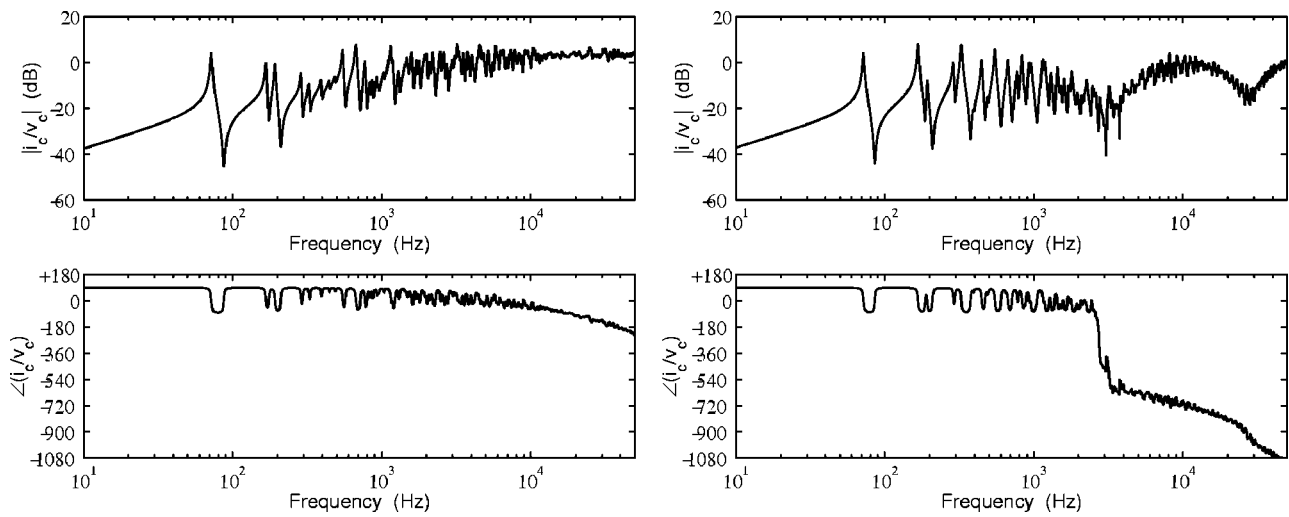


FIG. 11. Frequency response functions of a closely located velocity sensor and square piezoelectric patch actuator control unit (left, type a in Fig. 2) and a closely located velocity sensor and triangular piezoelectric patch actuator (right, type c in Fig. 2).

feedback control theory for SISO feedback control systems.^{3,7,47} The Bode and Nyquist plots of one sensor-actuator feedback loop frequency response function of the two control systems are shown in Figs. 11 and 12.

The left hand side amplitude plot in Fig. 11 highlights the typical feature of square strain actuators which more efficiently excite the panel at higher frequencies so that the amplitude of the sensor-actuator frequency response function grows as the frequency rises.⁵² In contrast the right hand side amplitude plot in Fig. 11 shows that the excitation generated by a triangularly shaped piezoelectric patch is modulated in frequency. This is probably due to a cancellation effect of the moment excitations generated along the lateral edges of the triangular patch in which case the actuation principally occurs via the transverse force generated at the tip of the triangular patch.

The left hand side phase plot of Fig. 11 indicates that the frequency response function generated by the square piezoelectric patch actuator with the velocity sensor at its center is

positive definite up to about 10 kHz. Thus this control system is only conditionally stable.^{3,7,47} Indeed, the Nyquist plot on the left hand side of Fig. 12 suggests that this control system would be unstable even with small feedback control gains since the higher-frequency part of the frequency response functions, with the larger amplitude, would encircle the stability point $-1 + j0$. The right hand side phase plot of Fig. 11 indicates that the frequency response function generated by the triangular piezoelectric patch actuator with the velocity sensor at its tip is positive definite only up to about 2.5 kHz where a sudden phase drop to about -540° occurs. Therefore this control system is also only conditionally stable. However in this case the Nyquist plot on the right hand side of Fig. 12 suggests that for this control system a relatively large gain margin is available since the left hand side of the plot is about five times smaller than the loops on the right hand side. This effect is due to the modulation of the excitation in frequency which combined with the complementary phase drops generates a Nyquist plot with

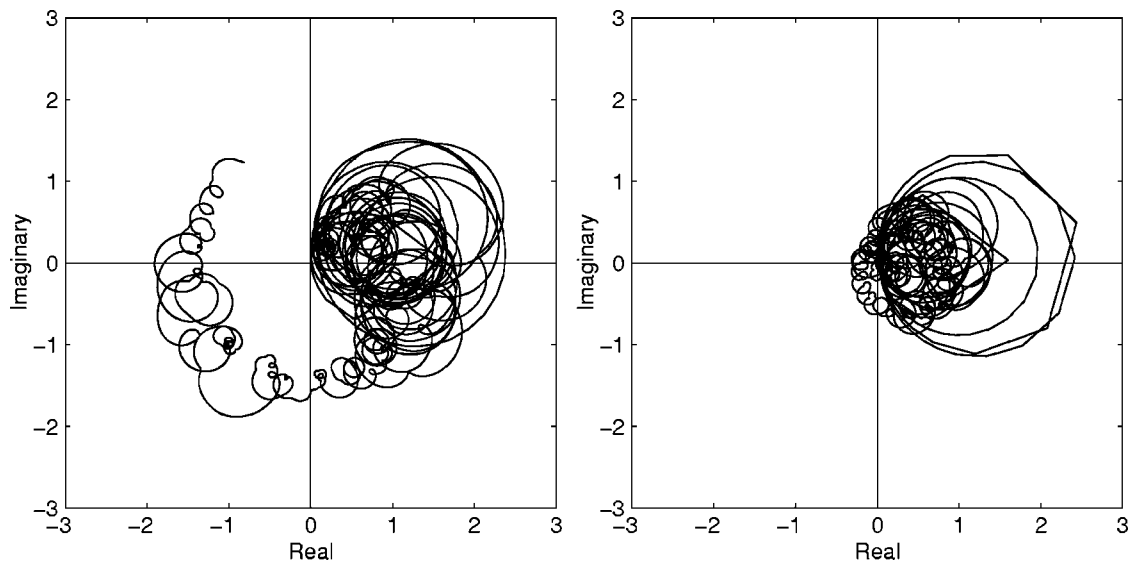


FIG. 12. Nyquist plots for the frequency response functions of a closely located velocity sensor and square piezoelectric patch actuator control unit (left, type a in Fig. 2) and a closely located velocity sensor and triangular piezoelectric patch actuator (right, type c in Fig. 2).

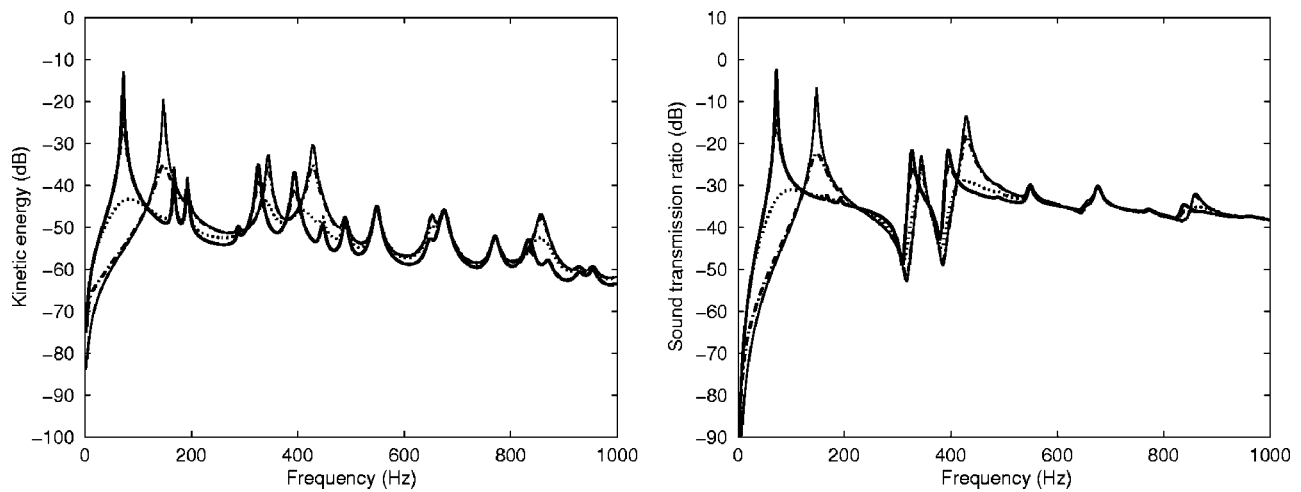


FIG. 13. Total kinetic energy (left) and sound transmission ratio (right) of the panel excited by a plane acoustic wave with no control (solid line) and with a 4×4 grid of closely located velocity sensors and square piezoelectric patches actuators (type b in Fig. 2) SISO feedback controller with feedback gains of 10 (dashed line), 10^2 (dotted line), 10^3 (dash-dotted line) and 10^4 (faint line).

the left hand side squeezed towards the imaginary axis. The modulated excitation of the triangularly shaped actuator depends on the angle of the lateral edges. A trial and error approach has been used in this study to find that the best stability effects are generated by a triangular patch with base $b = 40$ mm and height $a = 25$ mm.

The design of the closely located sensor-actuator control units is an important issue that has been briefly introduced in this section in order to contrast the intrinsic properties of sensor-actuator pairs made with either a square or a triangular strain actuator and an ideal velocity sensor. In practice the local dynamics effects of the sensor should also be taken into consideration as well as the mass and stiffening effects of the piezoelectric patch actuator. A detailed study of these issues is presented in Refs. 41, 53, 54 for a square piezoelectric patch actuator with a velocity sensor at its center.

V. SISO DIRECT VELOCITY FEEDBACK CONTROL

In this section the control effectiveness of the smart panels with SISO control units type (b) and type (d) in Fig. 2, which have either sixteen square piezoelectric patch actuators or sixteen triangular piezoelectric patch actuators driven by a single input signal, is investigated. In this case the stability of the two SISO control systems is fully analyzed by considering Bode and Nyquist plots of the sensor-actuator frequency response function.^{3,7,47}

A. Control effectiveness

The two plots in Fig. 13 show, respectively, the total kinetic energy and the sound transmission ratio of the smart panel with the SISO control system using the 4×4 array of square piezoelectric patch actuators at the center velocity sensors. As found with the MIMO control arrangement, the dashed and dotted lines in the two plots of Fig. 13 show that as the SISO control gain is raised up to an optimal control gain as the active damping generated by the feedback loop levels down most of the resonance peaks. However, in this case the SISO control system does not damp down the resonances due to natural modes of the panel which have no

volumetric vibration component since the sum of the sixteen sensors outputs goes to zero and thus the SISO control loop becomes ineffective. For example, no damping is introduced on the second, third and fourth resonances which are related to the (2,1), (1,2) and (2,2) natural modes of the panel. Beyond the optimal control gain, this trend is inverted and, as shown by the dash-dotted and faint lines, a new set of low-frequency resonances emerge whose amplitudes are similar, if not higher, than those of the panel with no control. Also in this case this phenomenon results from the control spillover effect where a large control gain tends to pin the smart panel at the control positions so that the response of the smart panel is rearranged into that of a lightly damped panel which is pinned at the sixteen control positions.

The two plots in Fig. 14 indicate that, as the control gain is raised as the normalized total kinetic energy and normalized sound transmission ratio, integrated from 0 Hz to 1 kHz, monotonically fall down and reach maximum reductions, respectively, of 12 dB and 8 dB. Therefore the SISO control arrangement is not able to replicate the 17 dB reduction of the total kinetic energy produced by the equivalent MIMO control system. In contrast it nearly generates the 9 dB reduction of the sound transmission ratio produced by the equivalent MIMO control system. This is due to the fact that the error signal used in the SISO feedback control loop is proportional to the volumetric vibration of the smart panel which generates most of the sound radiation at low frequency.^{12,16} When relatively high control gains are implemented then the pinning effect generated at the control positions brings the response of the smart panel back to the levels with no control while its sound radiation is increased by about 2.5 dB than in the case of no control. This is due to the increased sound radiation efficiency of the new resonant modes compared to that of the lower order modes of the unconstrained simply supported panel.¹

The two plots in Fig. 15 show the total kinetic energy and the sound transmission ratio of the smart panel with the SISO control system using the sixteen triangularly shaped piezoelectric patch actuators, with the base edges evenly dis-

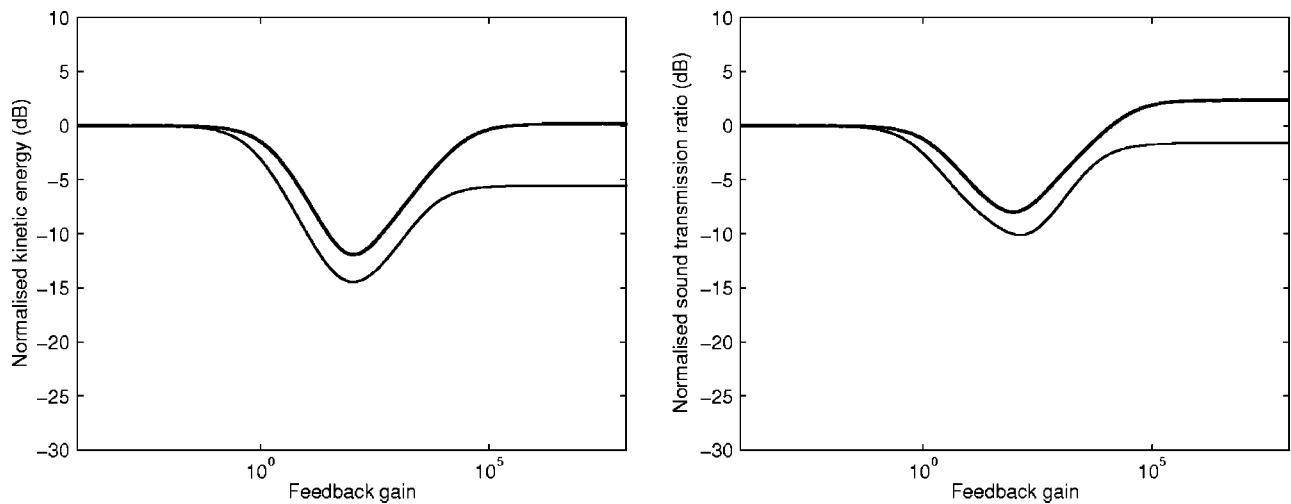


FIG. 14. Normalized total kinetic energy (left) and sound transmission ratio (right), integrated between 0 Hz and 1 kHz, plotted against the gain in the SISO velocity feedback controllers, h , for the 4×4 grid of closely located velocity sensors and square piezoelectric patches actuators (solid line, type b in Fig. 2) and the 16 closely located velocity sensors and triangular piezoelectric patches actuators (faint line, type d in Fig. 2).

tributed along the perimeter of the panel, and velocity sensors at the vertices opposite to the base edges. Comparing the dashed lines in these two plots with those of Fig. 13 it can be noticed that this SISO control system produces at the low-frequency resonances larger damping effects than the SISO control system using the sixteen square actuators. However it still does not generate damping effects at the resonance frequencies related to the natural modes of the panel with no volumetric component. Also, when large control gains, well above the optimal one, are implemented then the SISO control system with the sixteen triangular actuators generates new resonance frequencies some of which occurs at rather different frequencies than those obtained with the SISO control system using square actuators.

This type of behavior is confirmed by the faint lines of the two plots in Fig. 14 which indicates that the SISO arrangement with sixteen triangular actuators brings the maximum reduction of the normalized kinetic energy at 14.5 dB and the maximum reduction of the normalized sound trans-

mission ratio at 10 dB in comparison to the 12 dB and 8 dB, respectively, obtained with the SISO control system using the sixteen square actuators. However, as found for the SISO system with square actuators, the SISO arrangement with sixteen triangular actuators is not able to replicate the 19 dB reduction of the normalized kinetic energy obtained with the equivalent MIMO control system. In contrast it closely generates the 11 dB reduction of the normalized sound transmission ratio produced by the equivalent MIMO control system. Again this is due to the fact that the error signal used in the SISO feedback control loop is proportional to the volumetric vibration of the smart panel which generates most of the sound radiation at low frequency.^{12,16} When relatively high control gains are implemented, then the pinning effect at the sixteen control positions close to the edges of the panel produces a response of the smart panel which is about 5 dB lower than in the case of no control since the vibrating surface has been reduced to that delimited by the sixteen control positions. The frequency-averaged sound radiation of the

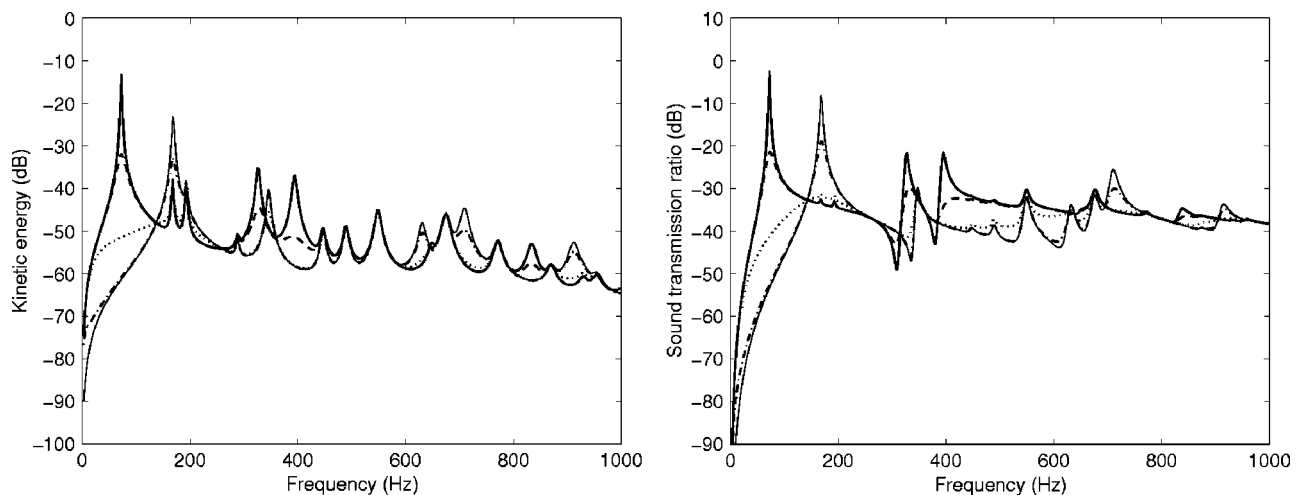


FIG. 15. Total kinetic energy (left) and sound transmission ratio (right) of the panel excited by a plane acoustic wave with no control (solid line) and with 16 closely located velocity sensors and 16 triangular piezoelectric patches actuators (type b in Fig. 2) SISO feedback controller with feedback gains of 10 (dashed line), 10^2 (dotted line), 10^3 (dash-dotted line) and 10^4 (faint line).

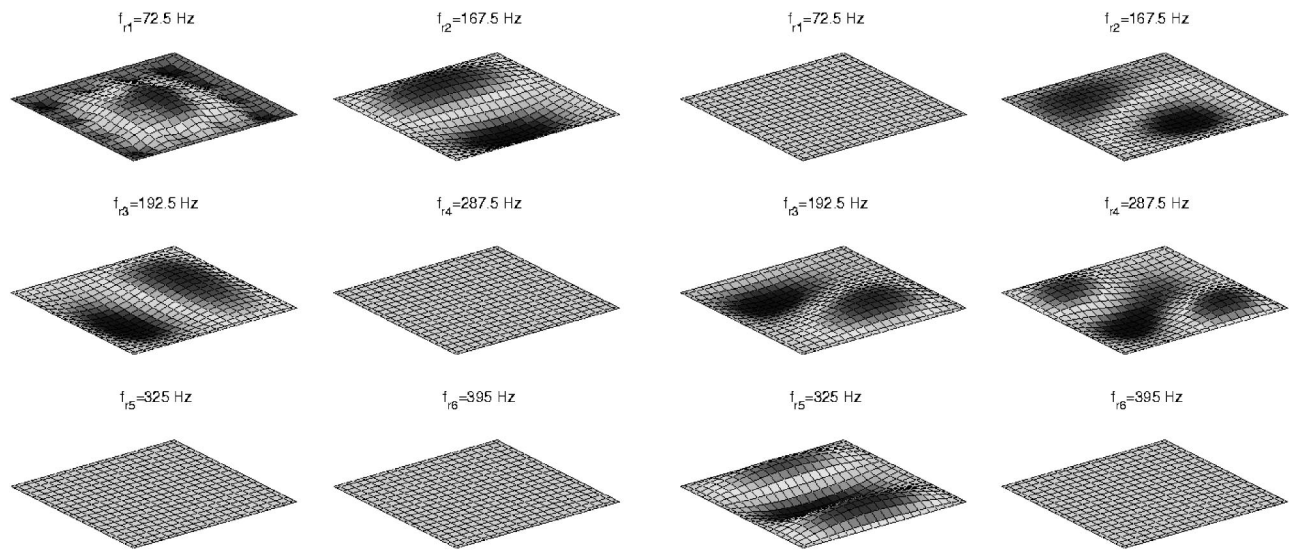


FIG. 16. Deflection shapes of the panel at the first six resonance frequencies when the optimal feedback control gain is implemented in the 4×4 grid of closely located velocity sensors and square piezoelectric patches actuators (left, type b in Fig. 2) and the 16 closely located velocity sensors and triangular piezoelectric patches actuators (right, type d in Fig. 2).

panel is instead reduced by just 2 dB, probably because the reduction of the response of the smart panel is balanced by the increased radiation efficiencies of the new resonant modes generated by the pinning effects along the perimeter of the smart panel.

Figures 16 and 17 show the deflections shapes of the panels in correspondence with the first six resonances when either the optimal control gain, that give the best control effects, or when a very large control gain is implemented in the smart panels with either the SISO control system made with sixteen square actuators (left hand side plots) or the SISO control system made with sixteen triangular actuators (right hand side plots). Also in this case, because the deflection shapes occurs over a wide range of amplitudes, the plots in Figs. 16 and 17 have been normalized to have the same

maximum deflections which are also equal to those in Figs. 8 to 10.

Comparing the two plots in Fig. 16 with that of Fig. 8 indicates that when the optimal control gains of the SISO control systems are implemented, then the response of the panel at some of the first six resonance frequencies is generally modified in such a way as the deflection shapes of the smart panel are not anymore controlled by the co-respective natural modes of the panel. This is due to the active damping action which, as seen for the MIMO control systems, tends to reduce the contribution of the resonant modes. However, as shown by the dashed lines on the left hand side plots in Figs. 13, the SISO control system with square actuators cannot damp down the resonances related to even-even or even-odd modes, such as those at 167.5 and 192.5 Hz relative to

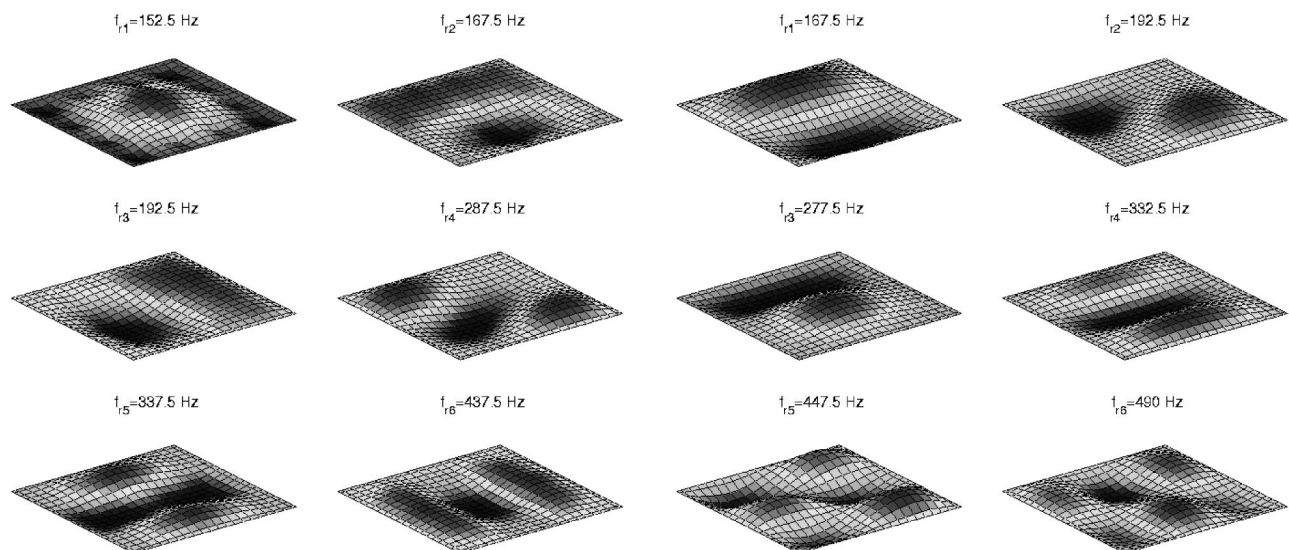


FIG. 17. Deflection shapes of the panel at the first six new resonance frequencies generated by a very large feedback gain in the 4×4 grid of closely located velocity sensors and square piezoelectric patches actuators (left, type b in Fig. 2) and the 16 closely located velocity sensors and triangular piezoelectric patches actuators (right, type d in Fig. 2).

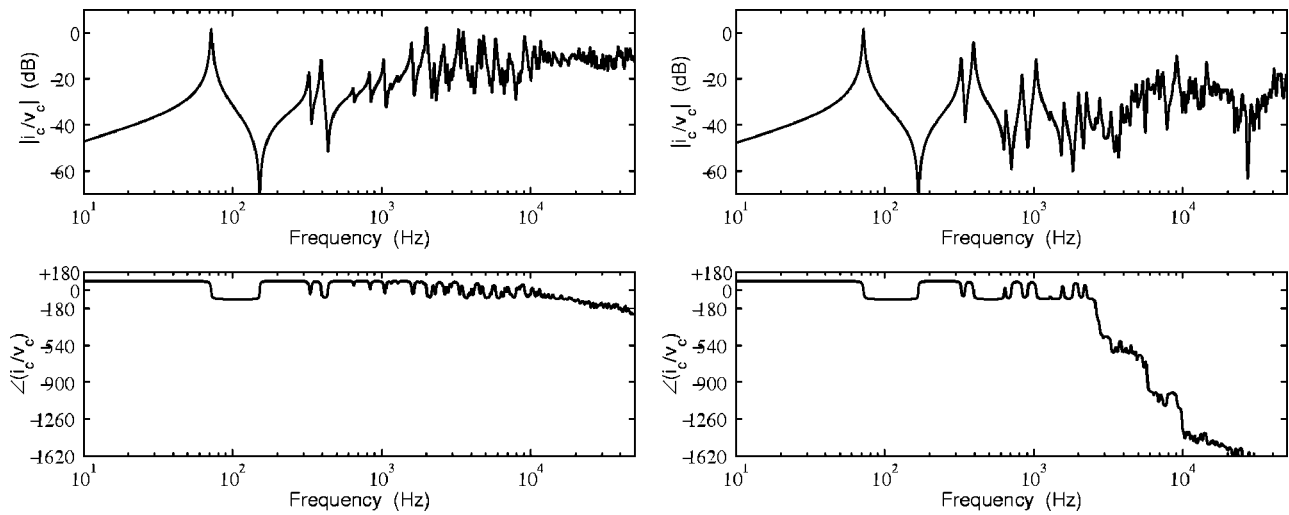


FIG. 18. Frequency response functions of the 4×4 grid of closely located velocity sensors and square piezoelectric patches actuators (left, type b in Fig. 12) and 16 closely located velocity sensors and triangular piezoelectric patches actuators (right, type d in Fig. 2).

the (2,1) and (1,2) modes, and thus the deflection shapes for these two resonance frequencies are exactly the same as those when there is no control. A similar phenomenon is found for the SISO system with triangular actuators arranged along the perimeter of the panel, even though in this case the deflection shapes relative to the resonances at 167.5 and 192.5 Hz are similar to (2,1) and (1,2) natural modes but rotated by an angle of 45° . This is probably due to the combined effects due to the azimuthal angle of the primary acoustic wave and the spacing of the error sensors along the perimeter of the panel. Finally the deflection shapes relative to the second three resonances generated by the system with square actuators and the deflection shapes relative to the first and sixth resonances generated by the system with triangular actuators do not show any type of vibration pattern. This is probably because, as shown by the dotted lines on the left hand side plots of Figs. 13 and 15, the optimal control gains tend to move up these resonance frequencies so that the plots in Fig. 16 show an off-resonance deflection shape.

As shown in the left hand side plot of Fig. 13, when very large control gains are implemented in the SISO control system with square actuators, a new set of resonance frequencies is generated. This is because, as shown on the left hand side plot of Fig. 17, the control actuators produce a 4×4 grid of pinning points that slightly modify the characteristic deflection shapes of the first six resonance frequencies of the unconstrained panel, which are shown in Fig. 8. The left hand side plot in Figs. 15 shows that when very large control gains are implemented in the SISO control system with triangular actuators the new set of resonance frequencies occurs at much higher frequencies. This is because the pinning effect generated by the control actuators generates higher order modes as one can deduce from the deflection shapes of the first six new resonances shown on the right hand side plot of Fig. 17.

B. Control stability

The stability properties of the two SISO control systems which, as shown in Figs. 2(b) and 2(d), are formed either by

a sixteen square piezoelectric patch actuator at the center velocity sensors or by sixteen triangularly shaped piezoelectric patch actuator with velocity sensors at the vertices are investigated using the classic feedback control theory for SISO feedback control systems.^{3,7,47} The Bode and Nyquist plots of the sensor-actuator open loop frequency response functions for the two control systems are therefore shown in Figs. 18 and 19. As found for a single control unit with a square actuator used in the decentralized MIMO control system, the left hand side amplitude plot in Fig. 18 highlights the typical rising trend due to the higher frequency excitation efficiency of square strain actuators⁵² although in this case there is a drop of the amplitude in the frequency range between 100 Hz and 1 kHz and the rising effect is less effective above about 10 kHz. This is probably due to the fact that between 100 Hz and 1 kHz and above 10 kHz the response of the panel is primarily controlled by structural modes with little or no volumetric vibration component so that the sum of the sensor outputs from the sixteen velocity sensors is relatively low. Also, as found for a single control unit with a triangular actuator used in the decentralized MIMO control system, the right hand side amplitude plot in Fig. 18 shows that the excitation generated by the sixteen triangularly shaped piezoelectric patches is modulated in frequency. Moreover in this case there is not an overall rising trend of the excitation so that the amplitude of the sensor-actuator frequency response function in correspondence to resonances below 1 kHz is at least equal to or much higher than those in correspondence to the higher-frequency resonances. This is probably due to a combination of effects where on one hand the excitation generated by the sixteen triangularly shaped actuators is reduced by local cancellation phenomena and on the other hand the sum of the sixteen control signals at the tips of the triangular actuators is also reduced by cancellation phenomena due to the contribution of higher order modes of the panel with a nonvolumetric vibration component.

The left hand side phase plot of Fig. 19 indicates that the frequency response function generated by the square piezoelectric patch actuators with the velocity sensor at their cen-

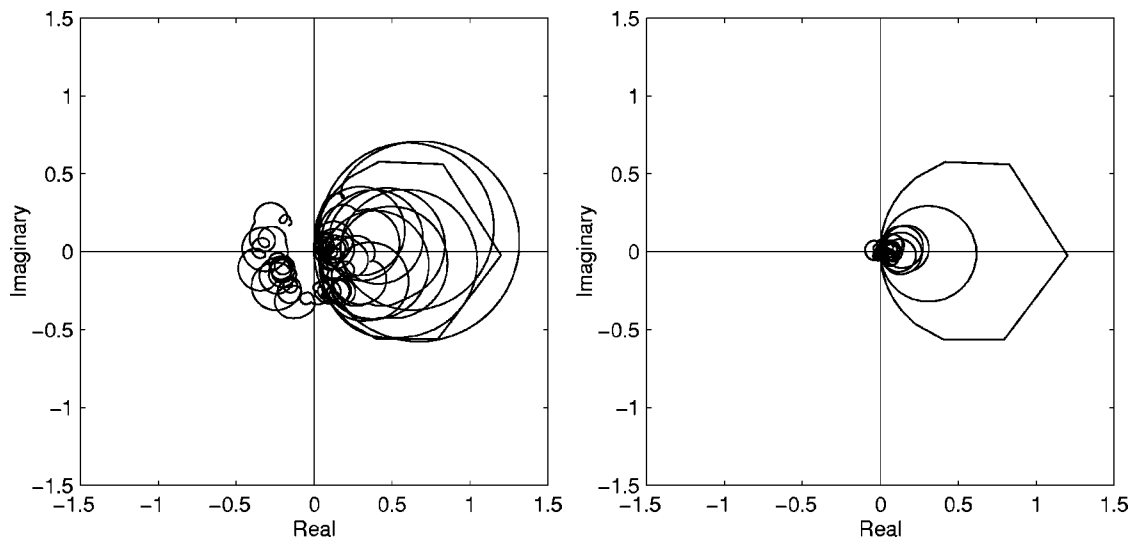


FIG. 19. Nyquist plots for the frequency response functions of the 4×4 grid of closely located velocity sensors and square piezoelectric patches actuators (left, type b in Fig. 2) and 16 closely located velocity sensors and triangular piezoelectric patches actuators (right, type d in Fig. 2).

ters is positive definite up to about 10 kHz so that this SISO control system is only conditionally stable.^{3,7,47} However, since above 10 kHz the amplitude of the frequency response function is about 10 dB lower than those of the first few resonant modes and that in the frequency band between 1 kHz and 10 kHz, the left hand side of the Nyquist plot on the left hand side of Fig. 19 is squeezed towards the imaginary axis so that a limited range of control gains could be implemented without the higher-frequency part of the frequency response function encircles the stability point $-1 + j0$. As found for a single control unit used in the MIMO control system, the right hand side phase plot of Fig. 18 indicates that the frequency response function generated by the sixteen triangular piezoelectric patch actuators with the velocity sensors at their tips is positive definite only up to about 2.5 kHz where a sudden phase drop to -540° occurs which also makes this SISO control system only conditionally stable. However, the fact that the amplitude of the sensor-actuator frequency response function above 1 kHz is modulated in frequency and relatively lower than at frequencies below 1 kHz gives the Nyquist plot shown on the right hand side of Fig. 19 which suggests that for this control system a relatively large gain margin is available. Indeed the left hand side of the Nyquist plot is about twenty times smaller than the loops on the right hand side so that large control gains could be implanted without generating instabilities. The small circles on the left hand side of the Nyquist plot are again due to the combination of periodic drops of the amplitude and co-respective drops of the phase of the sensor-actuator frequency response function.

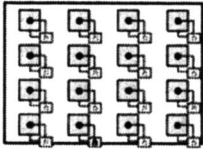
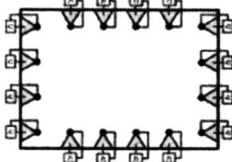
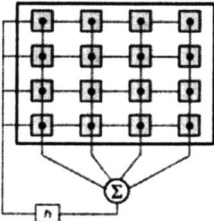
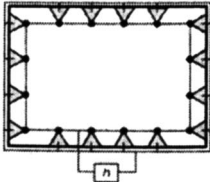
Therefore both SISO control system either with sixteen square control actuators and sixteen velocity sensors at their centers or with sixteen triangular control actuators and sixteen velocity sensors at the vertices could be used to implement direct velocity feedback control. In particular, the system with sixteen triangularly shaped actuators enables the implementation of relatively large control gains without causing instabilities.

VI. CONCLUSIONS

In this study the active structural acoustic control effectiveness of a new smart panel with sixteen triangularly shaped piezoelectric patch actuators having the base edges evenly distributed along the perimeter of the panel and velocity sensors at the vertices opposite to the base edges has been assessed and contrasted with that of a conventional smart panel made by a 4×4 array of square piezoelectric patch actuators evenly distributed over the surface of the panel at the centers velocity sensors. For both systems the control effectiveness and stability issues of MIMO decentralized or SISO velocity feedback control architectures have been analyzed.

As summarized in Table III the MIMO system with square actuators and velocity sensors at their centers relatively large reductions of the integrated kinetic energy and sound transmission ratio between 0 and 1 kHz can be achieved with a maximum value of, respectively, -17 and -9 dB. However the MIMO system with triangular actuators and velocity sensors at the tips produces even larger reductions, respectively, of -19 and -11 dB. Therefore the new system with triangular actuators and velocity sensors at the tips which are evenly distributed along the borders of the panel is not just able to replicate the results of the conventional smart panel with sixteen square piezoelectric patch actuators and velocity sensors at their centers. On the contrary it produces larger control reductions in particular with reference to the far field sound radiation. This is probably due to the fact that the active damping action is generated along the borders of the panel where indeed the sound radiation occurs at low frequencies below coincidence. The same type of behavior is obtained for the case where a SISO control architecture is implemented except that the reductions of the integrated kinetic energy and sound transmission ratio between 0 and 1 kHz go down, respectively, to -12 and -8 dB for the system with square actuators and velocity sensors at their centers and, respectively, to -14.5 and

TABLE III. Maximum frequency-averaged reductions of the total kinetic energy and sound transmission ratio with optimal feedback gains.

Smart panel	Change in kinetic energy dB	Change in sound transmission ratio dB
(a) MIMO-decentralised 	-17	-9
(c) MIMO-decentralised 	-19	-11
(b) SISO 	-12	-8
(d) SISO 	-14.5	-10

- 10 dB for the system with triangular actuators and velocity sensors at the tips. Since for both control arrangements the error sensors are evenly spaced either over the surface of the panel or along its perimeter, then no control signal is generated in correspondence to even-even or even-odd resonant modes which are therefore not controlled. Since the sound radiation efficiency of these modes is relatively low, then this phenomenon has little effect on the overall reduction of the sound radiation but, as confirmed by the data in Table III, it has relatively large effects on the vibration response of the panel and thus on its near field sound radiation.

The stability of one decentralized control unit for the two control arrangements has also been assessed by considering the Bode and Nyquist plots of the sensor-actuator frequency response function. This analysis has indicated that the triangular actuator is less effective at higher frequencies than the square actuator. Also, the triangular shaping generates a point force at the tip of the actuator where the point velocity sensor is placed so that a better collocation effect is obtained than with the square actuator. These two phenom-

ena makes the system with triangularly shaped actuators and velocity sensors at the tips to be relatively more stable so that direct velocity feedback control loops could be implemented up to relatively large control gains. Similar type of results have also been obtained for the SISO control architecture in which case the stability is even enhanced by the fact that the sensor does not measure the vibration contribution of even-even or even-odd resonant modes.

- ¹F. J. Fahy, *Sound and Structural Vibration* (Academic, London, 1985).
- ²C. R. Fuller, S. J. Elliott, and P. A. Nelson, *Active Control of Vibration* (Academic, London, 1996).
- ³R. L. Clark, W. R. Saunders, and G. P. Gibbs, *Adaptive Structures*, 1st ed. (Wiley, New York, 1998).
- ⁴S. J. Elliott and P. Gardonio, "Active vibroacoustic control," in *Responsive Systems For Active Vibration Control*, edited by A. Preumont (Kluwer Academic, London, 2002), pp. 59–82.
- ⁵P. Gardonio, "A review of active techniques for aerospace vibration and noise control," *J. Aircr.* **39**, 206–214 (2002).
- ⁶S. J. Elliott, *Signal Processing for Active Control*, 1st ed. (Academic, London, 2001).
- ⁷A. Preumont, *Vibration Control of Active Structures*, 2nd ed. (Kluwer Academic, London, 2002).
- ⁸J. Q. Sun, "Some observations on physical duality and collocation of structural control sensors and actuators," *J. Sound Vib.* **194**, 765–770 (1996).
- ⁹M. J. Balas, "Direct velocity control of large space structures," *J. Guid. Control* **2**, 252–253 (1979).
- ¹⁰G. V. Borgiotti, "The power radiated by a vibrating body in an acoustic fluid and its determination from boundary measurements," *J. Acoust. Soc. Am.* **88**, 1884–1893 (1990).
- ¹¹K. A. Cunefare, "The minimum multimodal radiation efficiency of baffled finite beams," *J. Acoust. Soc. Am.* **90**, 2521–2529 (1991).
- ¹²S. J. Elliott and M. E. Johnson, "Radiation modes and the active control of sound power," *J. Acoust. Soc. Am.* **94**, 2194–2204 (1993).
- ¹³W. T. Baumann, W. R. Saunders, and H. H. Robertshaw, "Active suppression of acoustic radiation from impulsively excited structures," *J. Acoust. Soc. Am.* **90**, 3202–3208 (1991).
- ¹⁴W. T. Baumann, F.-S. Ho, and H. H. Robertshaw, "Active structural acoustic control of broadband disturbances," *J. Acoust. Soc. Am.* **92**, 1998–2005 (1992).
- ¹⁵R. L. Clark and D. E. Cox, "Multi-variable structural acoustic control with static compensation," *J. Acoust. Soc. Am.* **102**, 2747–2756 (1996).
- ¹⁶M. E. Johnson and S. J. Elliott, "Active control of sound radiation using volume velocity cancellation," *J. Acoust. Soc. Am.* **98**, 2174–2186 (1995).
- ¹⁷S. D. Snyder, N. Tanaka, and Y. Kikushima, "The use of optimally shaped piezoceramic-electric film sensors in the active control of free field structural radiation. Part 2: feedback control," *ASME J. Vibr. Acoust.* **118**, 112–121 (1996).
- ¹⁸P. Gardonio, Y.-S. Lee, S. J. Elliott, and S. Debois, "Analysis and measurement of a matched volume velocity sensor and uniform force actuator for active structural acoustic control," *J. Acoust. Soc. Am.* **110**, 3025–3031 (2001).
- ¹⁹A. Francois, P. De Man, and A. Preumont, "Piezoceramic array sensing of volume displacement: a hardware demonstration," *J. Sound Vib.* **244**, 395–405 (2001).
- ²⁰B. Bingham, M. J. Atalla, and N. W. Hagood, "Comparison of structural-acoustic control designs on an active composite panel," *J. Sound Vib.* **244**, 761–778 (2001).
- ²¹P. De Man, A. Francois, and A. Preumont, "Robust feedback control of a baffled plate via open-loop optimization," *ASME J. Vibr. Acoust.* **124**, 154–157 (2002).
- ²²Y.-S. Lee, P. Gardonio, and S. J. Elliott, "Volume velocity vibration control of a smart panel using a uniform force actuator and an accelerometer array," *J. Smart Mater. Struct.* **11**, 863–873 (2002).
- ²³K. Henriouille and P. Sas, "Experimental validation of collocated PVDF volume velocity sensor/actuator pair," *J. Vibr. Acoust.* **124**, 154–157 (2002).
- ²⁴J. S. Vipperman and R. L. Clark, "Implications of using collocated strain-based transducers for output active structural acoustic control," *J. Acoust. Soc. Am.* **106**, 1392–1399 (1999).

- ²⁵L. Meirovitch and S. Thangjitham, "Control of sound radiation from submerged plates," *ASME J. Vibr. Acoust.* **88**, 402–407 (1990).
- ²⁶R. L. Clark and D. E. Cox, "Multi-variable structural acoustic control with static compensation," *J. Acoust. Soc. Am.* **102**, 2747–2756 (1997).
- ²⁷W. Dehandschutter, K. Henriouille, J. Swevers, and P. Sas, "Feedback control of broadband sound radiation using a volume velocity sensor," *Proceedings of ACTIVE 97*, Budapest, Hungary, 1997, pp. 979–992.
- ²⁸J. S. Vipperman and R. L. Clark, "Acoustic power suppression of a panel structure using H_∞ output feedback control," *J. Acoust. Soc. Am.* **105**, 219–225 (1999).
- ²⁹D. E. Cox, G. P. Gibbs, R. L. Clark, and J. S. Vipperman, "Experimental robust control of structural acoustic radiation," *ASME J. Vibr. Acoust.* **121**, 433–439 (1999).
- ³⁰R. L. Clark and D. E. Cox, "Experimental demonstration of a band-limited actuator/sensor selection strategy for structural acoustic control," *J. Acoust. Soc. Am.* **106**, 3407–3414 (1999).
- ³¹G. C. Smith and R. L. Clark, "Tradeoffs in design complexity—temporal versus spatial compensation," *J. Sound Vib.* **228**, 1182–1194 (1999).
- ³²J. S. Vipperman and R. L. Clark, "Multivariable feedback active structural acoustic control using adaptive piezoelectric sensor/actuators," *J. Acoust. Soc. Am.* **105**, 219–225 (1999).
- ³³G. P. Gibbs, R. L. Clark, D. E. Cox, and J. S. Vipperman, "Radiation modal expansion: application to active structural acoustic control," *J. Acoust. Soc. Am.* **107**, 332–339 (2000).
- ³⁴R. L. Clark and D. E. Cox, "Active control design for acoustic radiation using mixed-norm optimization," *J. Acoust. Soc. Am.* **108**, 1345–1348 (2000).
- ³⁵G. C. Smith and R. L. Clark, "A crude method of loop-shaping adaptive structures through optimum spatial compensator design," *J. Sound Vib.* **247**, 489–508 (2001).
- ³⁶K. Seto, M. Ren, and F. Doi, "Modeling and feedback structural acoustic control of a flexible plate," *ASME J. Vibr. Acoust.* **123**, 18–23 (2001).
- ³⁷W. Chang, S. V. Varadan, and V. K. Varadan, "Design of robust vibration controller for smart panel using finite element model," *ASME J. Vibr. Acoust.* **124**, 265–276 (2002).
- ³⁸B. Petitjean and I. Legrain, "Feedback controllers for active vibration suppression," *Jo. Struct. Control* **3**, 111–127 (1996).
- ³⁹S. J. Elliott, P. Gardonio, T. C. Sors, and M. J. Brennan, "Active vibro-acoustic control with multiple feedback loops," *J. Acoust. Soc. Am.* **111**, 908–915 (2001).
- ⁴⁰P. Gardonio, E. Bianchi, and S. J. Elliott, "Smart panel with multiple decentralised units for the control of sound transmission. Part I: theoretical predictions," *J. Sound Vib.* **274**, 163–192 (2004).
- ⁴¹P. Gardonio, E. Bianchi, and S. J. Elliott, "Smart panel with multiple decentralised units for the control of sound transmission. Part II: design of the decentralized control units," *J. Sound Vib.* **274**, 193–213 (2004).
- ⁴²P. Gardonio, E. Bianchi, and S. J. Elliott, "Smart panel with multiple decentralised units for the control of sound transmission. Part III: control system implementation," *J. Sound Vib.* **274**, 215–232 (2004).
- ⁴³J. M. Sullivan, J. E. Jr. Hubbard, and S. E. Burke, "Modeling approach for two-dimensional distributed transducers of arbitrary spatial distribution," *J. Acoust. Soc. Am.* **99**, 2965–2974 (1996).
- ⁴⁴P. Gardonio and S. J. Elliott, "Driving point and transfer mobility matrices for thin plates excited in flexure," ISVR Technical Report No 277, 1999.
- ⁴⁵B. T. Wang, C. R. Fuller, and E. K. Dimtriadis, "Active control of noise transmission through rectangular plates using multiple piezoelectric or point force actuators," *J. Acoust. Soc. Am.* **90**, 2820–2830 (1991).
- ⁴⁶C. K. Lee, "Theory of laminated piezoelectric plates for the design of distributed sensor/actuators. Part I: governing equations and reciprocal relationships," *J. Acoust. Soc. Am.* **87**, 1144–1158 (1990).
- ⁴⁷L. Meirovitch, *Dynamics and Control of Structures* (Wiley, New York, 1990).
- ⁴⁸S. M. Joshi, *Control of Large Flexible Space Structures* (Springer-Verlag, Berlin, 1989).
- ⁴⁹K. J. Aström and B. Wittenmark, *Adaptive Control*, 2nd ed. (Addison-Wesley, New York, 1995).
- ⁵⁰S. J. Elliott, M. Serrand, and P. Gardonio, "Feedback stability limits for active isolation systems with reactive and inertial actuators," *Trans. ASME, J. Vib. Acoust.* **123**, 250–261 (2001).
- ⁵¹P. Gardonio and S. J. Elliott, "Modal response of a beam with a sensor-actuator pair for the implementation of velocity feedback control," *J. Sound Vib.* (in press).
- ⁵²M. J. Brennan, M. J. Day, S. J. Elliott, and R. J. Pinnington "Piezoelectric actuators and sensors," *Proceedings of the Active Control of Vibration IUTAM Symposium*, 5–8 September 1994, University of Bath, UK, 1994.
- ⁵³M. Gavagni, P. Gardonio, and S. J. Elliott, "Theoretical study of a velometer sensor-piezoelectric patch actuator pair for direct velocity feedback control systems," ISVR Technical Memorandum, 2004.
- ⁵⁴A. Poy, "Modelling matched piezoelectric sensor/actuator pairs for feedback control," M.Sc. thesis, University of Southampton, 2003.

Reconstruction of transient acoustic radiation from a sphere

Sean F. Wu, Huancai Lu, and Manjit S. Bajwa

Department of Mechanical Engineering, Wayne State University, Detroit, Michigan 48202

(Received 13 February 2004; revision received 2 November 2004; accepted 5 November 2004)

Transient near-field acoustical holography (NAH) formulation is derived from the Helmholtz equation least squares (HELs) method to reconstruct acoustic radiation from a spherical surface subject to transient excitations in a free field. To facilitate derivations of temporal solutions, we make use of the Laplace transform and expansion in terms of the spherical Hankel functions and spherical harmonics, with their coefficients settled by solving a system of equations obtained by matching an assumed-form solution to the measured acoustic pressure. To derive a general form of solution for a temporal kernel, we replace the spherical Hankel functions and their derivatives by polynomials, recast infinite integrals in the inverse Laplace transform as contour integrals in a complex s -plane, and evaluate it via the residue theorem. The transient acoustic quantities anywhere including the source surface are then obtained by convoluting the temporal kernels with respect to the measured acoustic pressure. Numerical examples of reconstructing transient acoustic fields from explosively expanding, impulsively accelerating, and partially accelerating spheres, and that from a sphere subject to an arbitrarily time-dependent excitation are depicted. To illustrate the effectiveness of HELs-based transient NAH formulations, all input data are collected along an arbitrarily selected line segment and used to reconstruct transient acoustic quantities everywhere. [Work supported by NSF.] © 2005 Acoustical Society of America. [DOI: 10.1121/1.1841771]

PACS numbers: 43.40.Sk, 43.40.Rj, 43.60.Pt [EGW]

Pages: 2065–2077

I. INTRODUCTION

Analyses of transient acoustic radiation are often encountered in engineering applications because most vibrating structures are subject to impulsive or transient force excitations. In many cases, the transient excitations are unspecified and therefore the structural vibration responses are unknown, which makes the analysis of resultant acoustic field difficult. It is worth noticing that there are available numerous papers on predicting the transient acoustic^{1–10} and electromagnetic fields^{11–15} from finite objects subject to given excitations or boundary conditions. Of particular interest is the paper by Hansen¹⁶ on predicting time-domain acoustic radiation using a spherical near-field scanning around an arbitrary object. In other words, a single set of time-domain near-field measurements on a scanning sphere that encloses an arbitrarily shaped source can yield the time-domain acoustic field everywhere outside this scanning sphere.

However, the literature on the reconstruction of a transient acoustic field is very scarce. This scarceness is due to a lack of an effective methodology to deal with an inverse acoustic radiation problem in the time domain based on the limited information of acoustic pressure measurements.

One possible way of tackling transient acoustic radiation from an arbitrary object is to use the Kirchhoff–Helmholtz integral theory,¹⁷ which correlates the field acoustic pressure to surface acoustic pressure and normal surface velocity. These surface acoustic quantities are governed by the Kirchhoff–Helmholtz integral equation and must be solved before the field acoustic pressure can be calculated. For an arbitrary surface, this Kirchhoff–Helmholtz integral theory can only be implemented numerically through the boundary element method (BEM) by discretizing the surface into elements and calculating the acoustic quantities on each of the

discrete nodes. Since the acoustic quantities on these nodes have different emission times for an observer in any fixed coordinate system, the integrals are time dependent. Accordingly, one must discretize the integrals in both time and spatial domains, thus making the numerical computations extremely time consuming.⁵

An alternative in transient NAH is to reconstruct the acoustic quantities in the frequency domain first and take an inverse Fourier transform to retrieve the time domain signals. Wang was the first to reconstruct transient acoustic radiation from a vibrating object in this manner.¹⁸ In his Ph.D. dissertation, Wang demonstrated that a transient acoustic field could be reconstructed in the frequency domain using the Helmholtz equation least squares (HELs) method,¹⁹ and then taking an inverse Fourier transform to recover the time history of the normal surface velocity response. The infinite integral in inverse Fourier transform was calculated by direct numerical integration. Needless to say, numerical computations involved in this reconstruction process were intensive. Therefore, a direct calculation of an inverse Fourier transform to reconstruct transient acoustic radiation is not advisable.

Hald showed another pioneering work on reconstruction of temporal acoustic field via the so-called nonstationary spatial transformation of sound field (NS-STSF).²⁰ NS-STSF is based on time-domain holography (TDH) that processes the acoustic pressures measured by using a two-dimensional microphone array with the neighboring microphones separated by a half-wavelength. TDH “can be seen as a sequence of snapshots of instantaneous pressure over the array area, the time separation between subsequent snapshots being equal to the sampling interval in A/D conversion. Similarly, the output of TDH is a time sequence of snapshots of a

selected acoustic quantity in a calculation plane parallel to the measurement plane.”²⁰ Therefore, what we see is an acoustic pressure or intensity distribution in the frequency domain at some fixed instants over the recorded measurement time period. This is why it is called nonstationary STSF (NS-STSF), but not transient STSF. Note that NS-STSF is subject to the same restrictions as STSF. Namely, it is valid for reconstructing acoustic quantities on a planar surface in free space only, and allows for the projection of acoustic quantities along the microphones perpendicular to a measurement plane.

The objective of this paper is to derive transient NAH formulations using the HELS method²¹ to visualize acoustic waves traveling in both space and time based on a finite number of acoustic pressure measurements in free space. It is worth mentioning that HELS is valid for a arbitrarily shaped sources²² and has been used successfully to reconstruct time-harmonic acoustic radiation from vibrating structures in both exterior^{23,24} and interior regions.^{25,26}

The difference between Hansen’s work and the present one is that the former is a forward problem in which the cause (sound source) is given and its effect (sound field) is sought. On the other hand, the reconstruction of a transient acoustic field is a backward problem in which the effect is partially known via measurements, and the cause on a source surface is sought. Mathematically, a backward problem may be ill posed because the input data may be incomplete and inaccurate. Therefore, reconstruction is much more difficult to handle than prediction. A common approach to an ill-posed reconstruction problem is to regularize it by suppressing the small singular values that tend to inflate the errors embedded in the input and distort the reconstructed acoustic image.

For simplicity, our attention is focused on a spherical surface (the extension to a nonspherical surface is shown in a separate paper). In Sec. II we present HELS formulations for reconstructing transient acoustic quantities on a sphere in a free field. In Sec. III we derive general formulations for the temporal kernels. Stability of the solutions thus obtained is discussed in Sec. IV. Also shown are transient formulations via a spherical wave expansion with its coefficients determined by the orthogonality properties of the spherical harmonics. The advantages and disadvantages of these two approaches are discussed in Sec. V. In Sec. VI we demonstrate numerical examples of reconstructing the transient acoustic quantities produced by explosively expanding, impulsively accelerating, and impulsively accelerating baffled spheres, as well as those by a sphere subject to an arbitrarily time-dependent excitation. Conclusions are drawn in Sec. VII.

II. TRANSIENT HELS FORMULATIONS

Consider transient acoustic radiation from a sphere in a free field of density ρ_0 and speed of sound c . Assume that the sphere is excited at $t=t_0$ and prior to that the acoustic pressure at any point is zero, i.e., $p(\mathbf{x};t)=0$ at $t<t_0$. Our objective is to reconstruct the transient acoustic quantities everywhere, including the source surface, based on acoustic

pressure signals measured on a conformal surface Γ around the source. To this end, let us first define the Laplace transform pair for a function $f(\mathbf{x};t)$ as

$$F(\mathbf{x};s) = \int_0^{\infty} f(\mathbf{x};t) e^{-st} dt \quad \text{and} \\ f(\mathbf{x};t) = \frac{1}{i2\pi} \int_{\epsilon-i\infty}^{\epsilon+i\infty} F(\mathbf{x};s) e^{st} ds, \quad (1)$$

where s is a complex variable in the Laplace or s domain, ϵ is a real constant that is greater than the real parts of all singularities of $F(\mathbf{x};s)$. Note that the function $f(\mathbf{x};t)$ in Eq. (1) satisfies the condition $\int_0^{\infty} |f(\mathbf{x};t)| e^{-\sigma t} dt < \infty$ for some finite real value of σ , and ϵ is strategically placed to ensure that the real parts of all singularities of $F(\mathbf{x};s)$ fall on the left-half s plane so the result is bounded as $t \rightarrow \infty$.²⁷ Accordingly, we can express the acoustic pressure in the s domain as an expansion,²¹ which can be written under the spherical coordinates as

$$P(r, \theta, \phi; s) = \sum_{n=0}^N \sum_{l=-n}^n h_n(\beta) Y_n^l(\theta, \phi) C_{nl}(s), \quad (2)$$

where $\beta = isr/c$, $h_n(\beta)$ are the spherical Hankel functions of order n of the first kind, $Y_n^l(\theta, \phi)$ represent the spherical harmonics, and $C_{nl}(s)$ stands for the expansion coefficients. To determine the values of $C_{nl}(s)$, we can take acoustic pressure measurements over an enclosure conformal to the source at $\mathbf{x}^\Gamma \equiv \{x_m^\Gamma\}^T \in \Gamma$, $m=1$ to M , where $M > J$; here $J = (1+N)^2$ indicates the total number of expansion terms in (2). In a matrix form, this can be written as

$$\mathbf{P}(\mathbf{x}^\Gamma; s) = \mathbf{\Psi}(\mathbf{x}^\Gamma; s) \mathbf{C}(s), \quad (3)$$

where $\mathbf{P}(\mathbf{x}^\Gamma; s)$ and $\mathbf{C}(s)$ represent the column vectors that contain measured acoustic pressures and expansion coefficients, respectively, and $\mathbf{\Psi}(\mathbf{x}^\Gamma; s)$ is the matrix whose elements are given by

$$\Psi_{nl}(r_m^\Gamma, \theta_m^\Gamma, \phi_m^\Gamma; s) = h_n(\beta_m^\Gamma) Y_n^l(\theta_m^\Gamma, \phi_m^\Gamma).$$

The solution to Eq. (3) can be expressed as

$$\mathbf{C}(s) = \mathbf{\Psi}(\mathbf{x}^\Gamma; s)^\dagger \mathbf{P}(\mathbf{x}^\Gamma; s); \quad (4)$$

here $\mathbf{\Psi}(\mathbf{x}^\Gamma; s)^\dagger = [\mathbf{\Psi}(\mathbf{x}^\Gamma; s)^H \mathbf{\Psi}(\mathbf{x}^\Gamma; s)]^{-1} \mathbf{\Psi}(\mathbf{x}^\Gamma; s)^H$ denotes a pseudoinversion of the transfer matrix $\mathbf{\Psi}(\mathbf{x}^\Gamma; s)$ and superscript H implies a conjugate transpose. Once the $\mathbf{C}(s)$ are specified the acoustic pressure on the source surface $\mathbf{x}_0 \in S_0$ or anywhere else can be reconstructed,

$$\mathbf{P}(\mathbf{x}_0; s) = \mathbf{G}^p(\mathbf{x}_0 | \mathbf{x}^\Gamma; s) \mathbf{P}(\mathbf{x}^\Gamma; s), \quad (5)$$

where $\mathbf{G}^p(\mathbf{x}_0 | \mathbf{x}^\Gamma; s) = \mathbf{\Psi}(\mathbf{x}_0; s) \mathbf{\Psi}(\mathbf{x}^\Gamma; s)^\dagger$ is the transfer matrix for reconstructing acoustic pressure.

The normal surface velocity of a source can then be obtained via the Euler’s equation,

$$\mathbf{V}(\mathbf{x}_0; s) = \mathbf{G}^v(\mathbf{x}_0 | \mathbf{x}^\Gamma; s) \mathbf{P}(\mathbf{x}^\Gamma; s), \quad (6)$$

where $\mathbf{G}^v(\mathbf{x}_0 | \mathbf{x}^\Gamma; s) = \mathbf{\Phi}(\mathbf{x}_0; s) \mathbf{\Psi}(\mathbf{x}^\Gamma; s)^\dagger$ is the transfer matrix for reconstructing the normal surface velocity, $\mathbf{\Phi}(\mathbf{x}_0; s)$

$= -(\rho_0 s)^{-1} \partial \Psi(\mathbf{x}_0; s) / \partial r_0$, and ρ_0 is the density of the medium. Note that since the source is a sphere, the normal derivative becomes a derivative with respect to the radius r_0 .

The time-domain acoustic pressure and normal component of velocity at the j th point on the source surface $x_0 \equiv (r_0, \theta_0, \phi_0)$ is obtained by taking an inverse Laplace transform of (5) and (6), which is expressible as a convolution in a time domain:²⁸

$$p_j(x_0; t) = \sum_{m=1}^M g_{jm}^p(x_0|x_m^\Gamma; t) \circ p_m(x_m^\Gamma; t), \quad (7)$$

$$v_j(x_0; t) = \sum_{m=1}^M g_{jm}^v(x_0|x_m^\Gamma; t) \circ p_m(x_m^\Gamma; t), \quad (8)$$

where $x_m^\Gamma \in \Gamma$, $m=1$ to M , and the temporal kernels $g_{jm}^p(x_0|x_m^\Gamma; t)$ and $g_{jm}^v(x_0|x_m^\Gamma; t)$ are given by

$$g_{jm}^p(x_0|x_m^\Gamma; t) = \frac{1}{i2\pi} \int_{\epsilon-i\infty}^{\epsilon+i\infty} G_{jm}^p(x_0|x_m^\Gamma; s) e^{st} ds, \quad (9a)$$

$$g_{jm}^v(x_0|x_m^\Gamma; t) = \frac{1}{i2\pi} \int_{\epsilon-i\infty}^{\epsilon+i\infty} G_{jm}^v(x_0|x_m^\Gamma; s) e^{st} ds. \quad (9b)$$

It is emphasized that general closed-form solutions for the temporal kernels $g_{jm}^p(x_0|x_m^\Gamma; t)$ and $g_{jm}^v(x_0|x_m^\Gamma; t)$ cannot be found since measurement and reconstruction locations are arbitrary. The reason for having a flexible measurement scheme is because it is impractical to stipulate that acoustic pressure measurements be taken on a fixed configuration, say, a spherical surface.

III. TEMPORAL KERNELS

The integrands in Eqs. (9a) and (9b) are given by $G_{jm}^p(x_0|x_m^\Gamma; s) = \Psi_j(x_0; s) \Psi_j(x_m^\Gamma; s)^\dagger$ and $G_{jm}^v(x_0|x_m^\Gamma; s) = \Phi_j(x_0; s) \Psi_j(x_m^\Gamma; s)^\dagger$, respectively, where $\Psi_j(x; s) \equiv \Psi_{nl}(r, \theta, \phi; s) = h_n(\beta) Y_n^l(\theta, \phi)$ and $\Phi_j \equiv \Phi_{nl}(x; s) = -(\rho_0 s)^{-1} h_n(\beta) Y_n^l(\theta, \phi)$; here the indices j , n , and l are related together via $j = n^2 + n + l + 1$, with n increasing from 0 to N and l from $-n$ to n .

To facilitate evaluations of the temporal kernels, we recast the infinite integrals in Eqs. (9) into contour integrals:²⁹

$$g_{jm}^p(x_0|x_m^\Gamma; t) = \frac{1}{i2\pi} \oint_l G_{jm}^p(x_0|x_m^\Gamma; s) e^{st} ds, \quad (10a)$$

$$g_{jm}^v(x_0|x_m^\Gamma; t) = \frac{1}{i2\pi} \oint_l G_{jm}^v(x_0|x_m^\Gamma; s) e^{st} ds, \quad (10b)$$

where contour l consists of a straight line from $-i \operatorname{Im}(s)$ to $+i \operatorname{Im}(s)$ and a semicircle of radius $|s|$ on the left-half s -plane in counterclockwise direction. The contribution of this integral along the semicircle is identically zero as $|s| \rightarrow \infty$. This is obvious if we notice that the integrand in Eq. (10) contains the factor e^{st} that decays exponentially as $|s| \rightarrow \infty$ and $t > 0$.

The contour integrals in (10) can be evaluated using the residue theory.³⁰ To this end, we replace the spherical Hankel functions and their derivatives in $G_{jm}^p(x_0|x_m^\Gamma; s)$ and $G_{jm}^v(x_0|x_m^\Gamma; s)$ by polynomials of s :

$$h_n\left(\frac{ir}{c} s\right) = -e^{-rs/c} \sum_{q=0}^n \frac{(n+q)!}{i^n 2^q q!(n-q)!} \left(\frac{c}{r}\right)^{q+1} \frac{1}{s^{q+1}}, \quad (11a)$$

$$h_n'\left(\frac{ir}{c} s\right) = e^{-rs/c} \sum_{q=0}^n \frac{(n+q)! [(rs/c) + (q+1)]}{i^n 2^q q!(n-q)!} \times \left(\frac{c}{r}\right)^{q+1} \frac{1}{rs^{q+1}}. \quad (11b)$$

Substitute Eq. (11) into $G_{jm}^p(x_0|x_m^\Gamma; s)$ and $G_{jm}^v(x_0|x_m^\Gamma; s)$, use partial expansion to recast the polynomial of $[s - (s_{jm}^p)_q]^{-1}$, and sum them up in terms of the distinct real and pairs of complex poles. Doing so, we obtain

$$G_{jm}^p(x_0|x_m^\Gamma; s) = e^{[t_0 + (r_m^\Gamma - r_0)/c]s} \left[\sum_{q=0}^{Q_{jm}^p - 1} \frac{\Pi_{jm}^p(s_{jm}^p)_q}{s - (s_{jm}^p)_q} + \sum_{q=0}^{\hat{Q}_{jm}^p - 1} \frac{\hat{\Pi}_{jm}^p(\hat{s}_{jm}^p)_q}{s^2 - 2 \operatorname{Re}(\hat{s}_{jm}^p)_q s + |\hat{s}_{jm}^p|_q^2} \right], \quad (12a)$$

$$G_{jm}^v(x_0|x_m^\Gamma; s) = e^{[t_0 + (r_m^\Gamma - r_0)/c]s} \left[\sum_{q=0}^{Q_{jm}^v - 1} \frac{\Pi_{jm}^v(s_{jm}^v)_q}{s - (s_{jm}^v)_q} + \sum_{q=0}^{\hat{Q}_{jm}^v - 1} \frac{\hat{\Pi}_{jm}^v(\hat{s}_{jm}^v)_q}{s^2 - 2 \operatorname{Re}(\hat{s}_{jm}^v)_q s + |\hat{s}_{jm}^v|_q^2} \right], \quad (12b)$$

where $(s_{jm}^p)_q$ and $(\hat{s}_{jm}^p)_q$ denote the q th distinct real pole and q th pair of complex poles for reconstructing the surface acoustic pressure, respectively, $\Pi_{jm}^p(s_{jm}^p)_q$ and $\hat{\Pi}_{jm}^p(\hat{s}_{jm}^p)_q$ stand for the residues that correspond to $(s_{jm}^p)_q$ and $(\hat{s}_{jm}^p)_q$, respectively, Q_{jm}^p and \hat{Q}_{jm}^p are the total numbers of distinct real and pairs of complex poles for $G_{jm}^p(x_0|x_m^\Gamma; s)$. Similarly, $(s_{jm}^v)_q$ and $(\hat{s}_{jm}^v)_q$ indicate the q th distinct real and complex poles for reconstructing the normal surface velocity, respectively, $\Pi_{jm}^v(s_{jm}^v)_q$ and $\hat{\Pi}_{jm}^v(\hat{s}_{jm}^v)_q$ are the residues that correspond to $(s_{jm}^v)_q$ and $(\hat{s}_{jm}^v)_q$, respectively, and Q_{jm}^v and \hat{Q}_{jm}^v are the total numbers of distinct poles and pairs of complex poles for $G_{jm}^v(x_0|x_m^\Gamma; s)$. The general solutions for the temporal kernels can be obtained by multiplying sums of the residues by $i2\pi$:

$$g_{jm}^p(x_0|x_m^\Gamma; t) = \left[\sum_{q=0}^{\Theta_{jm}^p-1} \Pi_{jm}^p(s_{jm}^p)_q e^{-(s_{jm}^p)_q t} + \sum_{q=0}^{\hat{\Theta}_{jm}^p-1} \frac{\sin(\omega_{jm,q}^p t + \gamma_{jm,q}^p)}{\omega_{jm,q}^p} \times \hat{\Pi}_{jm}^p(\hat{s}_{jm}^p)_q e^{-\sigma_{jm,q}^p t} \right] H\left(t - t_0 - \frac{r_m^\Gamma - r_0}{c}\right), \quad (13)$$

$$g_{jm}^v(x_0|x_m^\Gamma; t) = \left[\sum_{q=0}^{\Theta_{jm}^v-1} \Pi_{jm}^v(s_{jm}^v)_q e^{-(s_{jm}^v)_q t} + \sum_{q=0}^{\hat{\Theta}_{jm}^v-1} \frac{\sin(\omega_{jm,q}^v t + \gamma_{jm,q}^v)}{\omega_{jm,q}^v} \times \hat{\Pi}_{jm}^v(\hat{s}_{jm}^v)_q e^{-\sigma_{jm,q}^v t} \right] H\left(t - t_0 - \frac{r_m^\Gamma - r_0}{c}\right),$$

where $H(t - t_0 - r_m^\Gamma - r_0/c)$ is the Heaviside step function, and ω_{jm} , σ_{jm} , and γ_{jm} are defined as

$$\omega_{jm,q}^p = \text{Im}(\hat{s}_{jm}^p)_q, \quad (14a)$$

$$\sigma_{jm,q}^p = \text{Re}(\hat{s}_{jm}^p)_q, \quad (14b)$$

$$\gamma_{jm,q}^p = \tan^{-1}[-\text{Im}(\hat{s}_{jm}^p)_q / \text{Re}(\hat{s}_{jm}^p)_q], \quad (14c)$$

$$\omega_{jm,q}^v = \text{Im}(\hat{s}_{jm}^v)_q, \quad (14d)$$

$$\sigma_{jm,q}^v = \text{Re}(\hat{s}_{jm}^v)_q, \quad (14e)$$

$$\gamma_{jm,q}^v = \tan^{-1}[-\text{Im}(\hat{s}_{jm}^v)_q / \text{Re}(\hat{s}_{jm}^v)_q]. \quad (14f)$$

Equation (13) illustrates that the temporal kernels act like weighting functions. The first terms in the square brackets of Eq. (13) resulting from the distinct real poles $(s_{jm}^p)_q$ and $(s_{jm}^v)_q$ decay exponentially in time. The second terms in the square brackets that result from distinct complex poles $(\hat{s}_{jm}^p)_q$ and $(\hat{s}_{jm}^v)_q$ are sinusoidal; however, their amplitudes decay exponentially in time and decrease monotonically in frequency. The rates of exponential decay are dictated by the real parts $\sigma_{jm,q}^p$ and $\sigma_{jm,q}^v$, and the angular frequencies are depicted by the imaginary parts $\omega_{jm,q}^p$ and $\omega_{jm,q}^v$ of $(\hat{s}_{jm}^p)_q$ and $(\hat{s}_{jm}^v)_q$. By convoluting the temporal kernels with measured acoustic pressures, we can reconstruct the transient acoustic waves that travel in space and time.

IV. STABILITY OF SOLUTIONS

As in any inverse problem, we may face an ill-posedness difficulty in reconstructing the temporal acoustic quantities.

In time-harmonic NAH, this ill-posedness problem is dealt with via regularization that smoothes the dependence of solution on the input data.³¹ A common approach is to employ Tikhonov regularization (TR)³² with a certain parameter choice method together with singular value decomposition (SVD). The objective of TR is to strike a balance between minimal residues and minimal solution norms, while SVD enables one to expand an acoustic quantity in terms of the principal acoustic modes and eliminate the evanescent waves that may drop below the ambient noise level and cause distortions in reconstruction.³³

In transient NAH regularization can be done conveniently by eliminating the poles that stay close to the imaginary axis. Since the contour integral is carried out in the left-half s plane, the real part of a distinct pole is negative and the amplitude of the corresponding temporal kernel decays exponentially in time. This guarantees a stability of a temporal solution. However, when a pole approaches the imaginary axis, its real part diminishes and the amplitude of corresponding temporal kernel decays very slowly. Under this condition, the errors embedded in the input data may remain finite and be transferred to the reconstructed acoustic quantities.

To circumvent this difficulty, we can design a low-pass filter that eliminates all poles whose real parts are larger than a tolerance ϵ . Accordingly, we can rewrite Eq. (12) as

$$G_{jm}^p(x_0|x_m^\Gamma; s) = e^{[t_0 + (r_m^\Gamma - r_0)/c]s} \left[\sum_{q=0}^{\Theta_{jm}^p-1} \frac{\Pi_{jm}^p(s_{jm}^p)_q}{s - (s_{jm}^p)_q} + \sum_{q=0}^{\hat{\Theta}_{jm}^p-1} \frac{\hat{\Pi}_{jm}^p(\hat{s}_{jm}^p)_q}{s^2 - 2 \text{Re}(\hat{s}_{jm}^p)_q s + |\hat{s}_{jm}^p|_q^2} \right], \quad (15a)$$

$$G_{jm}^v(x_0|x_m^\Gamma; s) = e^{[t_0 + (r_m^\Gamma - r_0)/c]s} \left[\sum_{q=0}^{\Theta_{jm}^v-1} \frac{\Pi_{jm}^v(s_{jm}^v)_q}{s - (s_{jm}^v)_q} + \sum_{q=0}^{\hat{\Theta}_{jm}^v-1} \frac{\hat{\Pi}_{jm}^v(\hat{s}_{jm}^v)_q}{s^2 - 2 \text{Re}(\hat{s}_{jm}^v)_q s + |\hat{s}_{jm}^v|_q^2} \right], \quad (15b)$$

where Θ_{jm}^p and $\hat{\Theta}_{jm}^p$ are total numbers of the principal poles for reconstructing surface acoustic pressure, Θ_{jm}^v and $\hat{\Theta}_{jm}^v$ are those of the principal poles for reconstructing normal surface velocity in the time domain, and the rest poles whose real parts are larger than ϵ are eliminated. Thus, this low-pass filter effectively expands an acoustic quantity in terms of the principal acoustic modes that remain bounded in the time domain, which is equivalent to what truncated SVD does for time-harmonic NAH applications. Accordingly, the regularized temporal kernels can be written as

$$\begin{aligned}
g_{jm}^p(x_0|x_m^\Gamma;t) &= \left[\sum_{q=0}^{\Theta_{jm}^p-1} \Pi_{jm}^p(s_{jm}^p)_q e^{-(s_{jm}^p)_q t} \right. \\
&\quad + \sum_{q=0}^{\hat{\Theta}_{jm}^p-1} \frac{\sin(\omega_{jm,q}^p t + \gamma_{jm,q}^p)}{\omega_{jm,q}^p} \\
&\quad \left. \times \hat{\Pi}_{jm}^p(\hat{s}_{jm}^p)_q e^{-\sigma_{jm,q}^p t} \right] H\left(t-t_0 - \frac{r_m^\Gamma - r_0}{c}\right), \quad (16) \\
g_{jm}^v(x_0|x_m^\Gamma;t) &= \left[\sum_{q=0}^{\Theta_{jm}^v-1} \Pi_{jm}^v(s_{jm}^v)_q e^{-(s_{jm}^v)_q t} \right. \\
&\quad + \sum_{q=0}^{\hat{\Theta}_{jm}^v-1} \frac{\sin(\omega_{jm,q}^v t + \gamma_{jm,q}^v)}{\omega_{jm,q}^v} \\
&\quad \left. \times \hat{\Pi}_{jm}^v(\hat{s}_{jm}^v)_q e^{-\sigma_{jm,q}^v t} \right] H\left(t-t_0 - \frac{r_m^\Gamma - r_0}{c}\right).
\end{aligned}$$

V. SPHERICAL WAVE EXPANSION

Since in this paper we focus on a spherical surface, it is possible to reconstruct transient acoustic quantities using the spherical wave expansion theory. Assume that the transient acoustic field is radiated from a sphere of radius $r=r_0$. Then we can write the acoustic pressure as²⁹

$$P(r, \theta, \phi; s) = \sum_{n=0}^{\infty} \sum_{l=-n}^n h_n(isr/c) Y_n^l(\theta, \phi) D_{nl}(s), \quad (17)$$

where the expansion coefficients $D(s)$ can be determined by multiplying both sides of Eq. (17) by a complex conjugate of the spherical harmonics $Y_n^l(\theta, \phi)^*$, integrating the resultant equation over a sphere of radius $r=r_m$ that enclose the source and on which the acoustic pressures are measured, and then using the orthogonality properties of the spherical

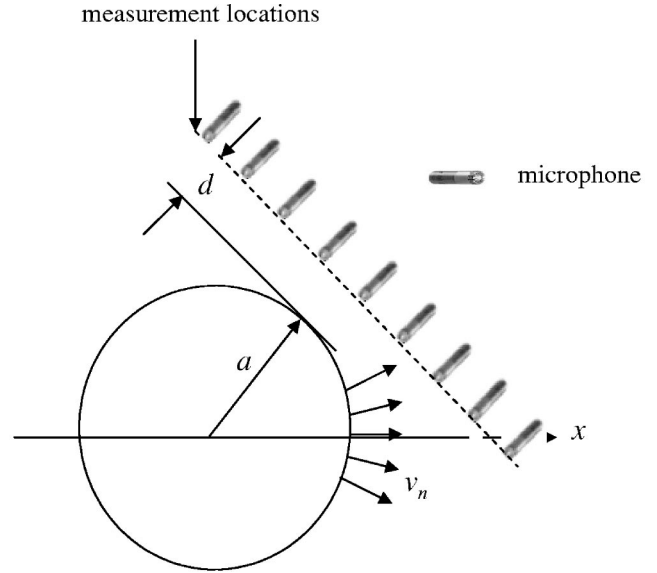


FIG. 1. Schematic of measurements of acoustic pressures from a partially accelerated sphere.

harmonics. Once $D(s)$ are specified, the surface acoustic pressure can be written as

$$\begin{aligned}
P(r_0, \theta, \phi; s) &= \sum_{n=0}^{\infty} \frac{h_n(isr_0/c)}{h_n(isr_m/c)} \sum_{l=-n}^n Y_n^l(\theta, \phi) \\
&\quad \times \int_{\Omega_m} \int P(r_m, \theta_m, \phi_m; s) \\
&\quad \times Y_n^l(\theta_m, \phi_m)^* d\Omega_m, \quad (18)
\end{aligned}$$

where a superscript $*$ implies a complex conjugate and $d\Omega_m = \sin \theta_m d\theta_m d\phi_m$ is the solid angle of integration over the measurement sphere.

The temporal surface acoustic pressure can now be obtained by taking an inverse Laplace transform of (18), which is expressible as a temporal convolution in the form of Eq. (7):

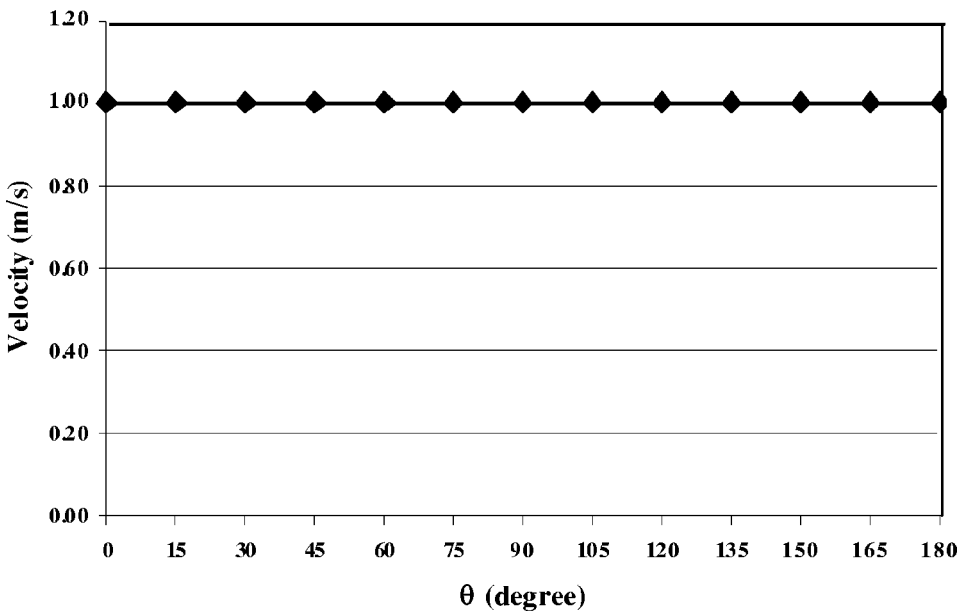
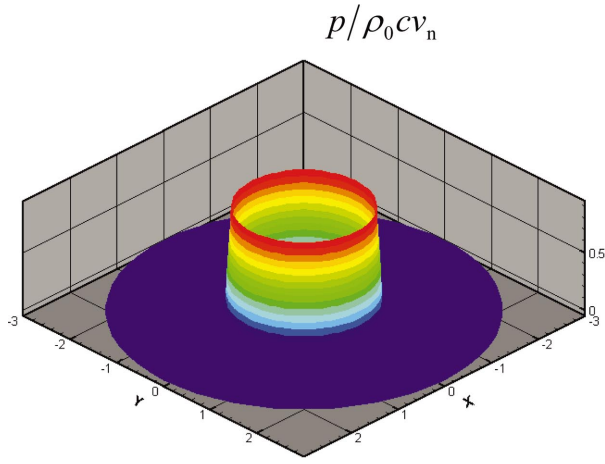
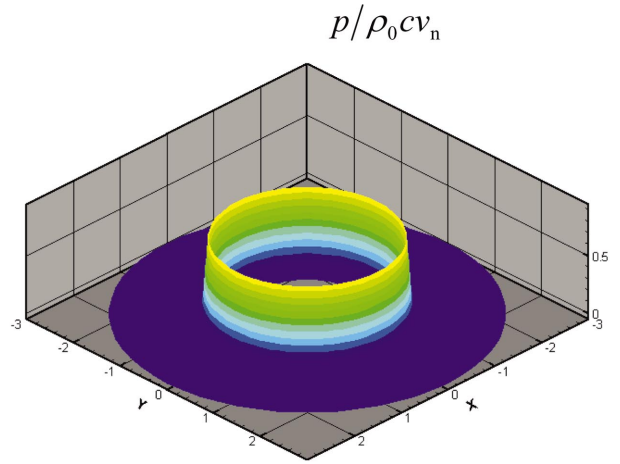


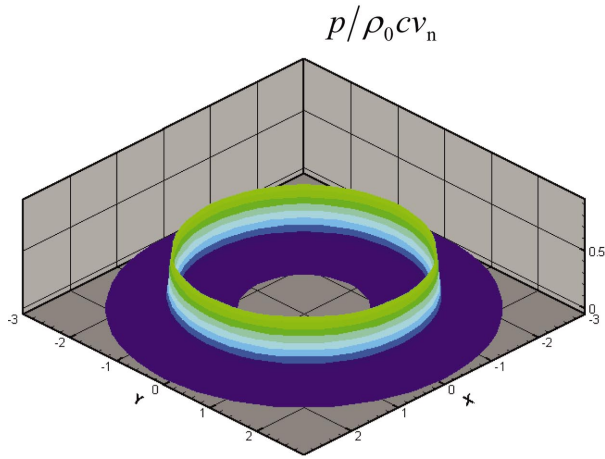
FIG. 2. A comparison of the reconstructed normal surface velocity distributions on the surface of an explosively expanding sphere and actual excitation. Solid line: excitation; \blacklozenge : reconstructed.



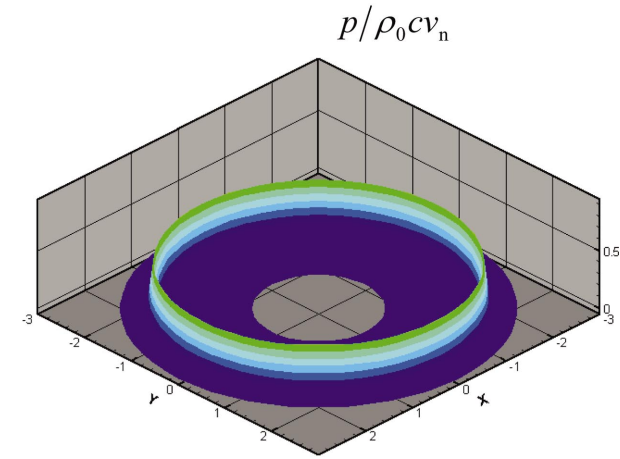
(3a) $t = 3.24$ (ms)



(3b) $t = 4.41$ (ms)



(3c) $t = 5.88$ (ms)



(3d) $t = 7.35$ (ms)



FIG. 3. Reconstructed 3-D transient acoustic pressure fields generated by an explosively expanding sphere.

$$p(r_0, \theta, \phi; t) = \sum_{n=0}^{\infty} A_n(r_0, r_m; t) \tilde{D}_n(r_m, \theta, \phi; t), \quad (19)$$

where A_n and \tilde{D}_n are defined, respectively, by

$$A_n(r_0, r_m; t) = \frac{1}{i2\pi} \int_{\epsilon - i\infty}^{\epsilon + i\infty} \frac{h_n(isr_0/c)}{h_n(isr_m/c)} e^{st} ds, \quad (20a)$$

$$\begin{aligned} \tilde{D}_n(r_m, \theta, \phi; t) &= \sum_{l=-n}^n Y_n^l(\theta, \phi) \\ &\times \int_{\Omega_m} p(r_m, \theta_m, \phi_m; t) \\ &\times Y_n^l(\theta_m, \phi_m)^* d\Omega_m. \end{aligned} \quad (20b)$$

The expansion coefficients A_n can be evaluated in exactly the same way as that for which the temporal kernel $g_{jm}^p(x_0|x_m^\Gamma; t)$ is obtained (see Sec. III). Thus, the spherical wave expansion is similar to HELS for a spherical surface; however, there are major differences between the two methods, as summarized below.

- (1) The spherical wave expansion is an infinite series and its solution converges uniformly in the region outside a sphere of radius $r = r_0$. HELS is a finite expansion and its solution is approximate, but nonetheless valid everywhere, even on an arbitrary source surface.²²
- (2) The spherical wave expansion uses the orthogonality property to determine its expansion coefficients, whose

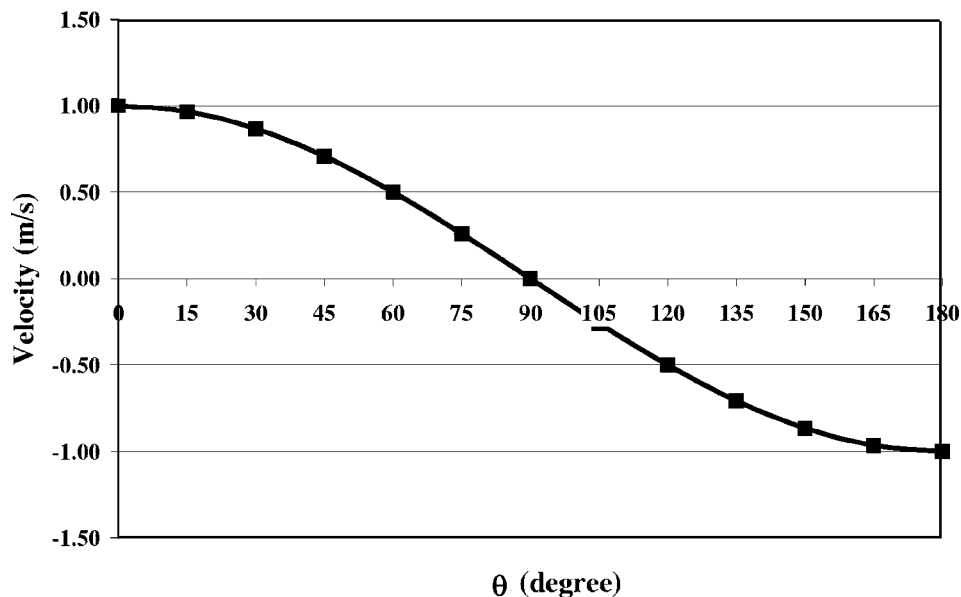


FIG. 4. A comparison of the reconstructed normal surface velocity distributions on the surface of an impulsively accelerated sphere and actual excitation. Solid line: excitation; ■: reconstructed.

values are fixed and independent of the number of expansion terms. HELS determines its coefficients by matching the assumed-form solution to the measured acoustic pressures and employs the least squares to minimize the residue norms.

- (3) The spherical wave expansion requires that the acoustic pressure measurements be taken over a spherical surface. HELS has no restrictions whatsoever on measurement locations.

Note that in practice it is not easy to take acoustic pressure measurements on a spherical surface. The flexibility of HELS in accommodating any measurement configuration can make it the method of choice for NAH. In Sec. VI, we illustrate numerical examples of reconstructing transient acoustic radiation from spheres by taking 12 measurements of the acoustic pressure along a straight line. Note that there is no way of knowing how many measurements are needed, because in general there is no prior knowledge of the source characteristics. If measurements are severely insufficient, the problem will be highly ill posed and the reconstructed image be greatly distorted. On the other hand, if too many measurements are taken, the reconstruction process can be too expensive and costly. So we must strike a balance between accuracy and cost in this case.

VI. NUMERICAL EXAMPLES

In this section, we use the transient NAH formulations (7) and (8) derived in Sec. III to reconstruct the transient acoustic quantities generated by a sphere. To facilitate validations of the reconstructed field, we consider a sphere of radius $r_0 = 1$ m subject to an impulsive excitation for which the analytic solution is readily available:

$$v(r_0, \theta_0, \phi_0; t) = \hat{v}(r_0, \theta_0, \phi_0)H(t - t_0), \quad (21)$$

where $\hat{v}(r_0, \theta_0, \phi_0)$ is the amplitude of the normal velocity on the source surface.

In time-harmonic NAH it has been shown that satisfactory reconstruction can be obtained by taking acoustic pres-

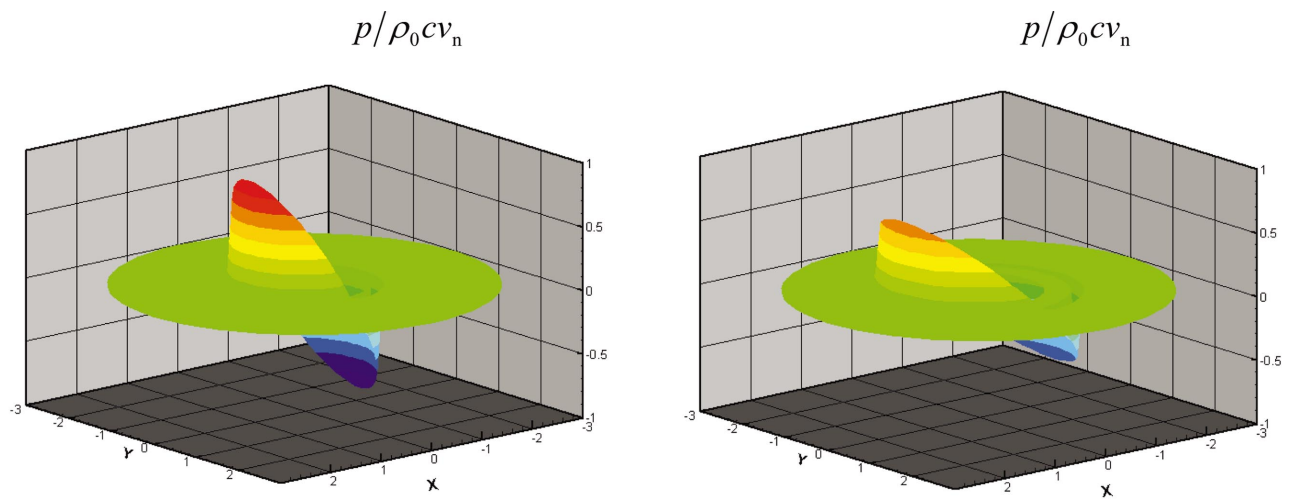
sure measurements on a surface conformal to the source geometry. This is especially true in transient NAH for the microphones can capture transient signal simultaneously. To test the effectiveness of the present formulations, we deliberately align the microphones along a line segment because this is easy to implement in practice. Figure 1 illustrates 11 measurement microphones that are placed on a straight line that intersects the x axis at 135° with a minimum distance of $d = 0.1$ m. Apparently, under such a measurement scheme, an acoustic signal cannot reach all microphones simultaneously, which makes reconstruction a challenge. Accordingly, we take a shot segment of measurements of acoustic pressures and ensure that all microphones have received the signals, although the amplitudes at these microphones may be quite different. The signals thus obtained are taken as input data to Eqs. (7) and (8) to reconstruct the entire acoustic field including the source surface in the time domain.

A. Explosively expanding sphere

The first example is concerned with reconstruction of transient acoustic radiation from an explosively expanding sphere in free space. Since the resultant field is spherically symmetric, the acoustic pressure is a function of the radial distance and time only. The analytic solution for an explosively expanding sphere is given by²⁹

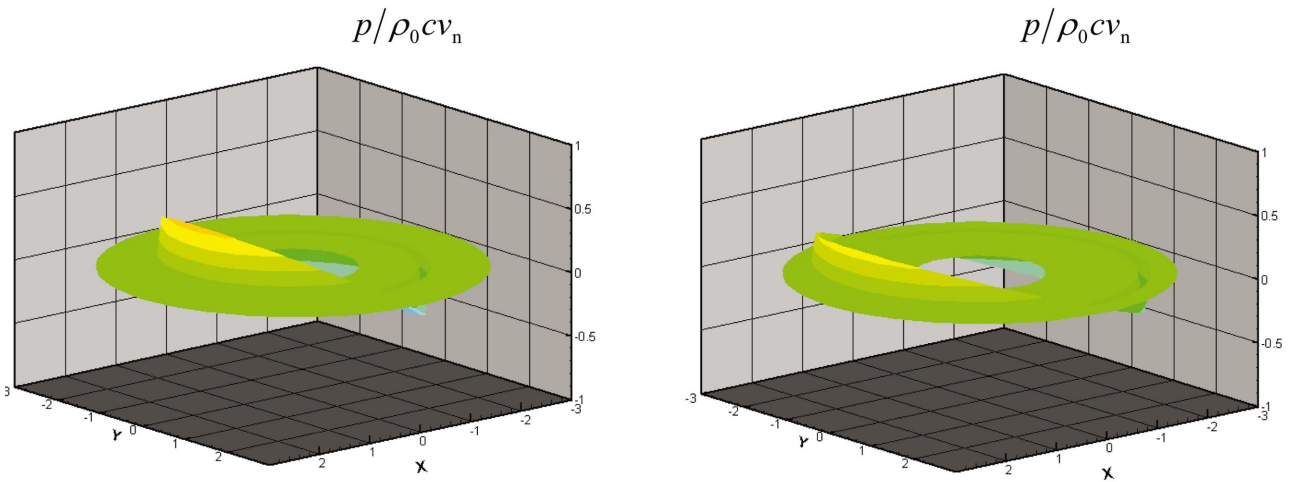
TABLE I. Comparison of coefficients ζ_j for a partially impulsively accelerating baffled sphere.

ζ_j	Reconstructed values	Analytic Values
ζ_1	$+6.69999E-02$	$+6.70000E-02$
ζ_2	$+1.87500E-01$	$+1.87500E-01$
ζ_3	$+2.70633E-01$	$+2.70633E-01$
ζ_4	$+3.00783E-01$	$+3.00781E-01$
ζ_5	$+2.74016E-01$	$+2.74016E-01$
ζ_6	$+1.98730E-01$	$+1.98730E-01$
ζ_7	$+9.34531E-02$	$+9.34529E-02$
ζ_8	$-1.76239E-02$	$-1.76239E-02$
ζ_9	$-1.10301E-01$	$-1.10301E-01$
ζ_{10}	$-1.65869E-01$	$-1.65869E-01$
ζ_{11}	$-1.75139E-01$	$-1.75140E-01$



(5a) $t = 3.24$ (ms)

(5b) $t = 4.41$ (ms)



(5c) $t = 5.88$ (ms)

(5d) $t = 7.35$ (ms)



FIG. 5. Reconstructed 3-D transient acoustic pressure fields generated by an impulsively accelerated sphere.

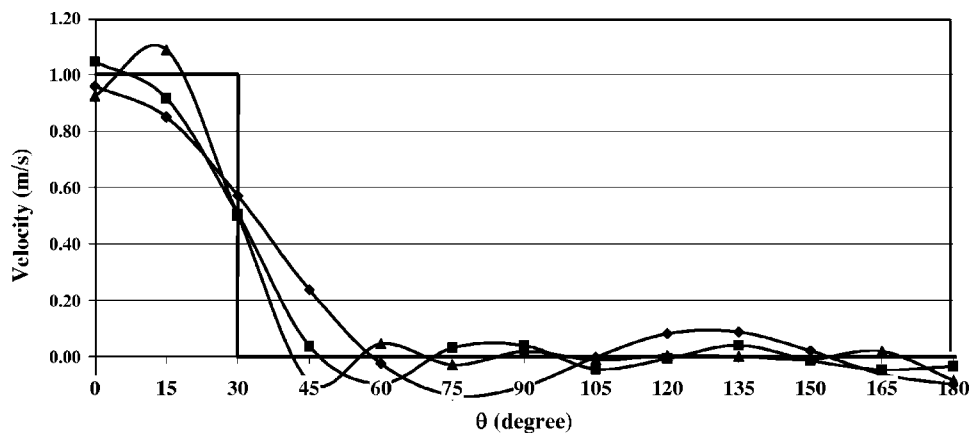


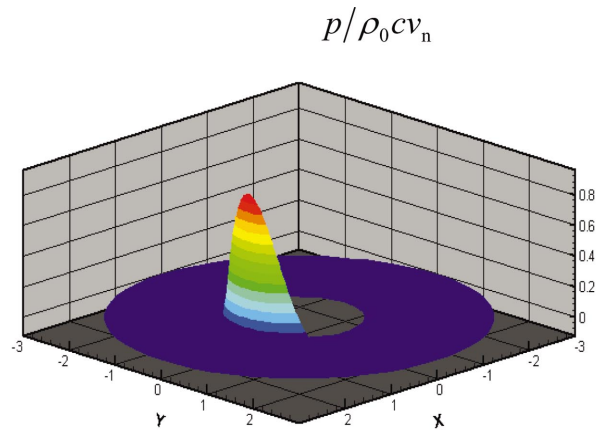
FIG. 6. A comparison of reconstructed normal surface velocity distributions on the surface of an impulsively accelerated baffled sphere and excitation. Solid line: excitation; \blacklozenge : reconstructed with $J=4$; \blacksquare : reconstructed with $J=8$; \blacktriangle : reconstructed with $J=11$.

TABLE II. Distinct poles for an impulsively accelerated baffled sphere.

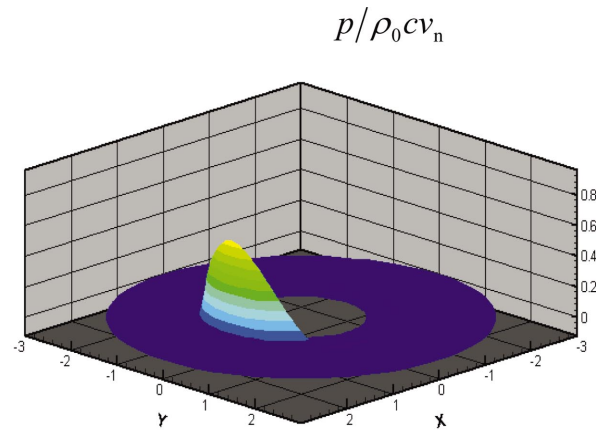
No.	Distinct Poles s_j
1	$-5.53363E+02 -i3.28058E+03$
2	$-1.41101E+03 -i2.46182E+03$
3	$-1.87779E+03 -i1.80707E+03$
4	$-2.16515E+03 -i1.19157E+03$
5	$-2.32467E+03 -i5.92509E+02$
6	$-2.37603E+03 +i0.00000E+00$
7	$-2.32467E+03 +i5.29509E+02$
8	$-2.16515E+03 +i1.19157E+03$
9	$-1.87779E+03 +i1.80707E+03$
10	$-1.41101E+03 +i2.46182E+03$
11	$-5.53363E+02 +i3.28058E+03$

$$\frac{p(r;t)}{\rho_0 c v} = \frac{r_0 e^{-(ct-(r-r_0))/r_0}}{r} H\left(t - \frac{(r-r_0)}{c}\right). \quad (22)$$

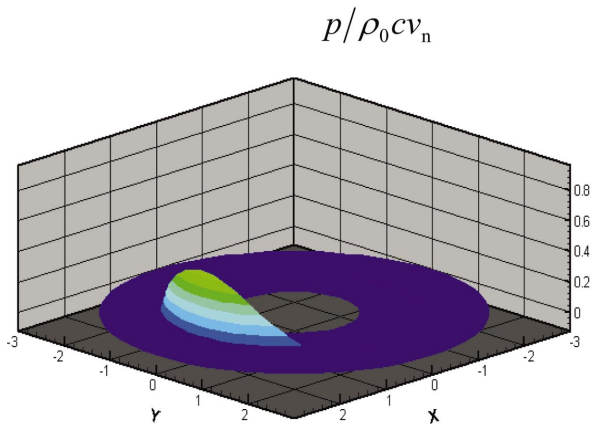
Equation (22) is used to generate transient acoustic pressures at the microphone locations. It is emphasized that the signal level at each microphone is different because the impulse cannot reach all microphones simultaneously. These acoustic pressure signals are taken as input to Eqs. (7) and (8) to reconstruct the transient acoustic field. In this case, there is only one distinct pole at $s = -c/r_0$, whose residue is r_0/r^Γ ; here r^Γ is the radial distance of a microphone to the center of the explosively expanding sphere ($r_0 < r^\Gamma$). The



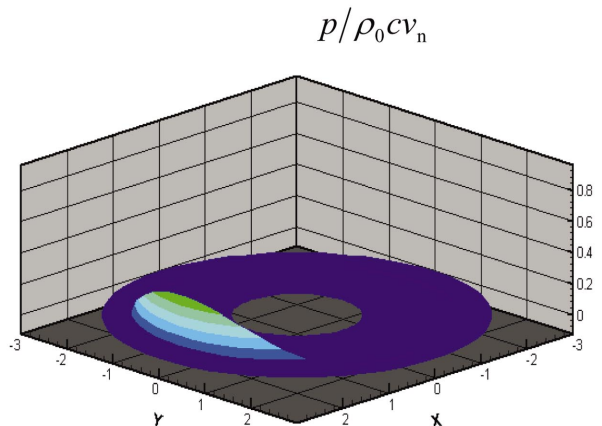
(7a) $t = 3.24$ (ms)



(7b) $t = 4.41$ (ms)



(7c) $t = 5.88$ (ms)



(7d) $t = 7.35$ (ms)



FIG. 7. Reconstructed transient acoustic pressure field from an impulsively accelerated baffled sphere.

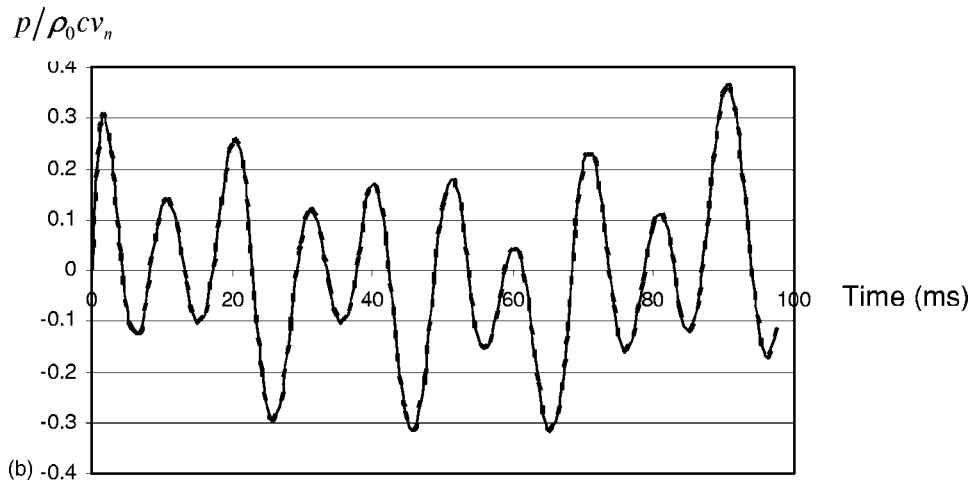
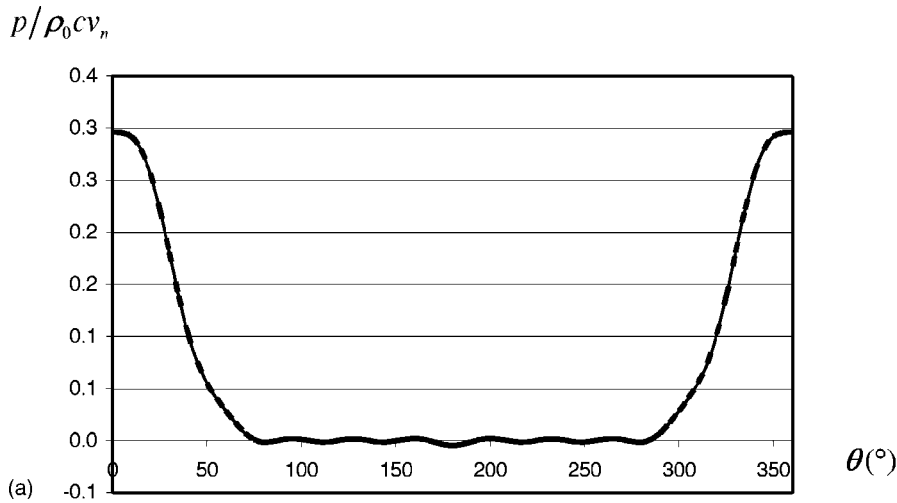


FIG. 8. A comparison of exact dimensionless acoustic pressure (solid line) and reconstructed ones (dashed line) on the surface of a partially vibrating sphere under an arbitrarily time-dependent excitation. (a) Distribution at $t = 2.529$ ms with $e = 5.875 \times 10^{-9}$; (b) time history at $\theta = 20^\circ$ with $e = 4.602 \times 10^{-8}$, where $e = \|\{p_{\text{rec}}\} - \{p_{\text{exact}}\}\|_2 / \|\{p_{\text{exact}}\}\|_2$.

test results show that solutions thus obtained converge rapidly. This is no surprise since the input data are exact. Hence perfect reconstruction of surface acoustic pressure and normal surface velocity can be obtained directly.

For brevity, we only exhibit the comparison of reconstructed normal surface velocities with the actual excitation (see Fig. 2) and three-dimensional images of acoustic pressure fields at arbitrarily selected time instances t

$t = 3.24$ (ms), 4.41 (ms), 5.88 (ms), and 7.35 (ms) (see Fig. 3).

B. Impulsively accelerated sphere

The second example is reconstruction of transient acoustic radiation from an impulsively accelerated sphere in free space, whose analytic solution is given by²⁹

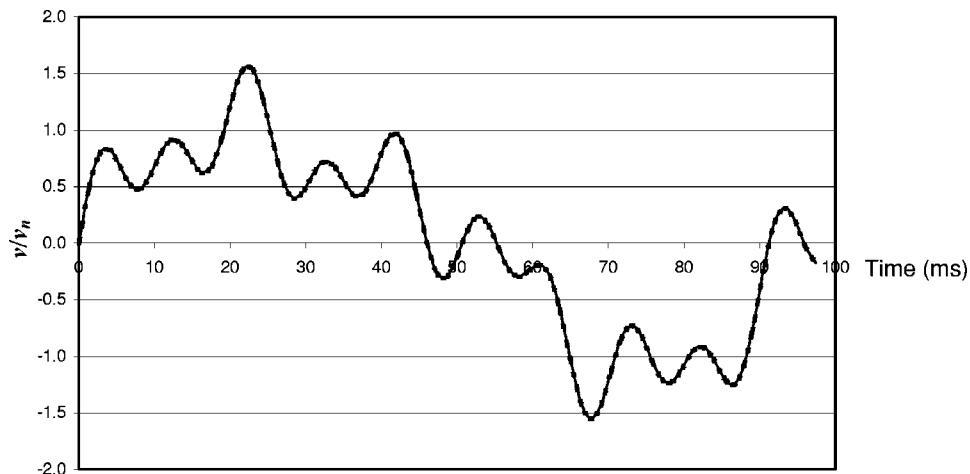
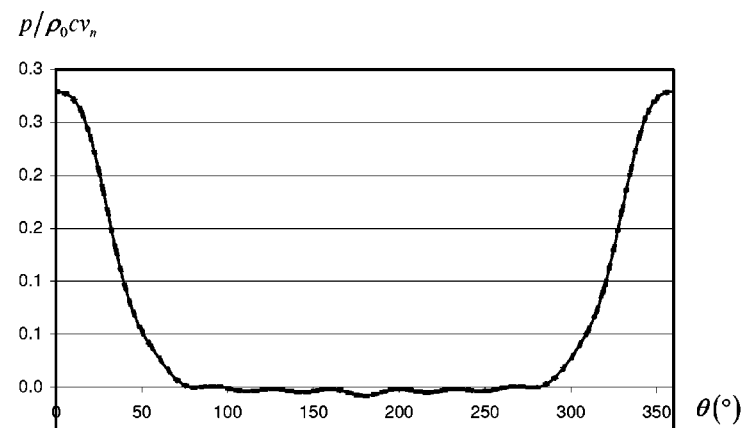
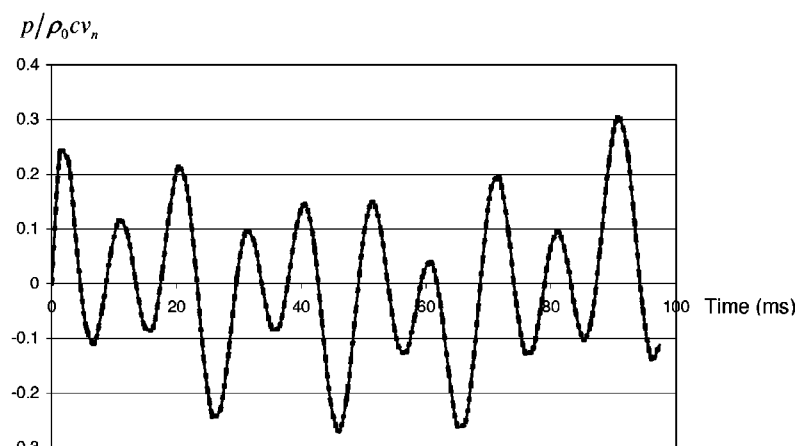


FIG. 9. A comparison of the exact normal velocity history (solid line) and reconstructed one (dashed line) on the surface of a partially vibrating sphere under an arbitrarily time-dependent excitation at $\theta = 20^\circ$ with $e = 6.262 \times 10^{-3}$, where $e = \|\{v_{\text{rec}}\} - \{v_{\text{exact}}\}\|_2 / \|\{v_{\text{exact}}\}\|_2$.



(a)



(b)

$$\frac{p(r, \theta; t)}{\rho_0 c v} = \frac{r_0 \cos \theta}{r} e^{-(ct - (r - r_0))/r_0} H\left(t - \frac{(r - r_0)}{c}\right) \times \left[\cos\left(\frac{ct - (r - r_0)}{r_0}\right) - \left(1 - \frac{r_0}{r}\right) \sin\left(\frac{ct - (r - r_0)}{r_0}\right) \right], \quad (23)$$

where $\cos \theta = \mathbf{n} \cdot \mathbf{e}_R$; here \mathbf{n} is the unit normal on the source surface and \mathbf{e}_R is the unit vector in the direction of wave propagation from the source to a receiver.

Equation (23) is used to generate the field acoustic pressures in the same manner as that described in the preceding section. These signals are substituted in Eqs. (7) and (8) to reconstruct the temporal acoustic quantities. In this case, there are a pair of complex conjugate poles located at $s = (-1 \pm i)(c/r_0)$. The corresponding residues are $(r_0/2r^\Gamma)[1 \mp (r_0/r^\Gamma - 1)i]$.

Figure 4 displays the reconstructed normal component of velocity, which agrees perfectly with the actual excitation. Figure 5 depicts the reconstructed acoustic pressure distributions produced by an impulsively accelerated sphere in free space at $t = 3.24$ (ms), 4.41 (ms), 5.88 (ms), and 7.35 (ms).

C. Impulsively accelerated baffled sphere

In this example, a small portion of the surface is impulsively accelerated and the rest area is rigidly baffled. Mathematically, we can write the normal surface velocity as

$$v(r_0, \theta, t) = \hat{v}_0(r_0, \theta) [H(\theta + \theta_0) - H(\theta - \theta_0)] H(t - t_0), \quad (24a)$$

where $2\theta_0$ is the vertex angle (in this example we select $2\theta_0 = 60^\circ$) and $\hat{v}(r_0, \theta)$ is the amplitude of the normal surface velocity. The normal surface velocity can be rewritten as

$$v(r_0, \theta, t) = \sum_{j=0}^{\infty} \zeta_j(r_0, \theta_0) P_j(\cos \theta) H(t - t_0), \quad (24b)$$

where $P_j(\cos \theta_0)$ represents the Legendre functions²⁹ and $\zeta_j(r_0, \theta_0)$ are the expansion coefficients. The analytic solution of transient acoustic pressure can be written as an infinite series.³⁰

Table I compares the reconstructed expansion coefficients with the exact values. Because of a rich content of high spatial frequencies around the corners at $\theta_0 = \pm 30^\circ$, as it is indicated by the Heaviside step function in Eq. (24a), a large number of expansion functions are required to approximate a sudden rise of the normal surface velocity. Figure 6 displays a comparison of the reconstructed normal surface velocity using different numbers of expansion functions with the actual excitation. Results indicate that the reconstructed normal surface velocity converges to the actual one as the number of expansion terms increases. Theoretically, if input data are complete and error free, the reconstructed normal surface velocity may converge to the true value as the number of expansion terms approaches infinity. However, this will not happen since in reality the input data are incomplete

FIG. 10. A comparison of exact dimensionless acoustic pressure (solid line) and reconstructed ones (dashed line) in the field produced by a partially vibrating sphere under an arbitrarily time-dependent excitation. (a) Distribution at $r = 1.1$ m, $t = 2.529$ ms, and $e = 3.125 \times 10^{-8}$; (b) Time history at $r = 1.1$ m, $\theta = 20^\circ$, and $e = 1.607 \times 10^{-8}$, where $e = \|\{p_{\text{rec}}\} - \{p_{\text{exact}}\}\|_2 / \|\{p_{\text{exact}}\}\|_2$.

and inaccurate. Therefore, regularization is needed to truncate the expansion or eliminate unstable poles so that errors in input are not amplified in reconstruction.

Table II lists the distinct poles s_j , $j=1$ to 11, in this case. Once these poles are identified and the corresponding residues are obtained, the temporal acoustic pressure can be reconstructed by the temporal kernel with measured acoustic pressures. Figure 7 shows reconstructed 3-D transient acoustic pressure fields produced by an impulsively accelerated baffled sphere at $t=3.24$ (ms), 4.41 (ms), 5.88 (ms), and 7.35 (ms).

D. A partially vibrating sphere under an arbitrarily time-dependent excitation

Finally, we show reconstruction of transient acoustic radiation from a partially vibrating sphere subject to arbitrarily time-dependent excitations in a free field. Without loss of generality, we express the normal surface velocity as a sum of harmonic functions with randomly selected amplitudes, frequencies, and phases,

$$v(r_0, \theta, t) = \hat{v}_0(r_0, \theta)[H(\theta + \theta_0) - H(\theta - \theta_0)] \times \sum_{j=1}^4 A_j \sin(2\pi f_j t + \vartheta_j) H(t - t_0), \quad (25)$$

where $A_1=1$, $A_2=0.2$, $A_3=0.3$, $A_4=0.3$, $f_1=0.5$, $f_2=23$, $f_3=56$, $f_4=100$, $\vartheta_1=10^\circ$, $\vartheta_2=40^\circ$, $\vartheta_3=30^\circ$, and $\vartheta_4=40^\circ$. The input data are generated by superimposing contributions from a partially vibrating sphere under individual harmonic excitation. To test the effectiveness of the formulations derived, we insert 5% spatially uncorrelated Gaussian noise to the input data and substitute them in Eqs. (7) and (8) to reconstruct the temporal acoustic quantities. To ensure a satisfactory reconstruction, we regularize the solutions by eliminating the poles that are close to the imaginary axis of the s plane.

Note that in this case the distinct poles remain unchanged (see Table II) since they depend on source geometry and measurement configuration. However, the residues are different because they are not only source geometry and measurement configuration dependent, but also excitation dependent. Once the residues are obtained, the temporal surface acoustic pressure and normal surface velocity are obtained by convoluting the temporal kernels over input acoustic pressures. As before, excellent reconstruction is obtained. For brevity, we illustrate the comparisons of the reconstructed acoustic pressure and normal velocity on the source surface in the time domain (see Figs. 8 and 9), the reconstructed acoustic pressure on a spherical surface of radius $r=1.1$ m at $t=2.529$ (ms) versus the benchmark values [see Fig. 10(a)], and the time history of acoustic pressure at a field point $r=1.1$ m and $\theta=0^\circ$ versus the benchmark values [see Fig. 10(b)].

VII. CONCLUSIONS

HELs formulations are extended to the reconstruction of transient acoustic radiation from a spherical source subject to arbitrarily time-dependent excitations in a free field. It is

shown that HELs enables one to obtain the satisfactory reconstruction of 3-D images of transient acoustic waves as they travel in space and time based on a few measurements taken along a finite line segment. Since HELs imposes no restrictions whatsoever on measurement locations and since the nature of HELs formulations is naturally suited for transient NAH, it can become the method of choice for visualizing transient acoustic quantities.

ACKNOWLEDGMENT

The work was supported by NSF Grant No. CMS-0245587.

- ¹J. Brillouin, "Rayonnement transitoire des sources sonores et problemes connexes," *Ann. Telecommun.* **5**, 160–172, 179–194 (1950).
- ²H. Huang, "Transient interaction of plane acoustic waves with a spherical elastic shell," *J. Acoust. Soc. Am.* **45**, 661–670 (1969).
- ³H. Huang, "Transient response of two fluid-coupled spherical elastic shells to an incident pressure pulse," *J. Acoust. Soc. Am.* **65**, 881–887 (1979).
- ⁴P. M. Pinsky and N. N. Abboud, "Transient finite element analysis of the exterior structural acoustics problem," *ASME J. Vib. Acoust.* **112**, 245–256 (1990).
- ⁵S. F. Wu, "Transient sound radiation from impulsively accelerated bodies," *J. Acoust. Soc. Am.* **94**, 542–553 (1993).
- ⁶H.-W. Chen and P. Stepanishen, "Acoustic transient radiation from fluid-loaded shells of revolution using time-dependent in vacuo eigenvector expansion," *J. Acoust. Soc. Am.* **95**, 601–616 (1994).
- ⁷P. Stepanishen and H.-W. Chen, "Acoustic time-dependent loading on elastic shells of revolution using the internal source density and singular value decomposition method," *J. Acoust. Soc. Am.* **99**, 1913–1923 (1996).
- ⁸L. L. Thompson and R. Huan, "Computation of transient radiation in semi-infinite regions based on exact nonreflecting boundary conditions and mixed time integration," *J. Acoust. Soc. Am.* **106**, 3095–3108 (1999).
- ⁹R. Huan and L. L. Thompson, "Accurate radiation boundary conditions for the time dependent wave equation on unbounded domains," *Int. J. Numer. Methods Eng.* **47**, 1569–1603 (2000).
- ¹⁰J. A. Hamilton and R. J. Astley, "Exact solutions for transient spherical radiation," *J. Acoust. Soc. Am.* **109**, 1848–1858 (2001).
- ¹¹W. C. Davidon, "Time-dependent multipole analysis," *J. Phys. A* **6**, 1635–1646 (1973).
- ¹²A. D. Yaghjian and T. B. Hansen, "Time-domain far fields," *J. Appl. Phys.* **79**, 2822–2930 (1996).
- ¹³E. Heyman, "Time-dependent plane-wave spectrum representations for radiation from volume source distributions," *J. Math. Phys.* **37**, 658–681 (1996).
- ¹⁴E. Heyman and A. J. Devaney, "Time-dependent multipoles and their application for radiation from volume source distributions," *J. Math. Phys.* **37**, 682–692 (1996).
- ¹⁵E. A. Marengo and A. J. Devaney, "Time-dependent plane wave and multipole expansions of the electromagnetic field," *J. Math. Phys.* **39**, 3643–3660 (1998).
- ¹⁶T. B. Hansen, "Spherical expansions of time-domain acoustic fields: Application to near-field scanning," *J. Acoust. Soc. Am.* **98**, 1204–1215 (1995).
- ¹⁷A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (McGraw-Hill, New York, 1981), Chap. 4, pp. 165–194.
- ¹⁸Z. Wang, "Helmholtz Equation-Least-Squares (HELs) method for inverse acoustic radiation problems," Ph.D. dissertation, Wayne State University, Detroit, Michigan, 1995.
- ¹⁹Z. Wang and S. F. Wu, "Helmholtz equation-least-squares method for reconstructing the acoustic pressure field," *J. Acoust. Soc. Am.* **102**, 2020–2032 (1997).
- ²⁰J. Hald, "Use of non-stationary STSF for the analysis of transient engine noise radiation," Brüel & Kjær, Skodsborgvej 307, DK-2850 Nærum, Denmark, 1999.
- ²¹S. F. Wu, "On reconstruction of acoustic pressure fields using the Helmholtz equation least squares method," *J. Acoust. Soc. Am.* **107**, 2511–2522 (2000).

- ²²V. Isakov and S. F. Wu, "On theory and applications of the Helmholtz equation least squares method in inverse acoustics," *Inverse Probl.* **18**, 1147–1159 (2002).
- ²³N. E. Rayess and S. F. Wu, "Experimental validations of the HELS method for reconstructing acoustic radiation from a complex vibrating structure," *J. Acoust. Soc. Am.* **107**, 2955–2964 (2000).
- ²⁴S. F. Wu, N. Rayess, and X. Zhao, "Visualization of acoustic radiation from a vibrating bowling ball," *J. Acoust. Soc. Am.* **109**, 2771–2779 (2001).
- ²⁵S. F. Wu and J. Yu, "Reconstructing interior acoustic pressure fields via Helmholtz equation least-squares method," *J. Acoust. Soc. Am.* **104**, 2054–2060 (1998).
- ²⁶M. Moondra and S. F. Wu, "Visualization of sound transmission into a vehicle passenger compartment," *Proceedings of the ASME Noise Control and Acoustic Division, NCA-Vol. 3* (CD ROM), IMECE2001/NCA-23540, November 2001.
- ²⁷B. C. Kuo, *Automatic Control System*, 5th ed. (Prentice–Hall, Englewood Cliffs, NJ, 1987), pp. 22–24.
- ²⁸E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic Press, San Diego, 1999).
- ²⁹P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, NJ, 1986).
- ³⁰M. C. Junger and D. Feit, *Sound, Structures, and Their Interactions* (MIT Press, Cambridge, MA, 1972).
- ³¹E. G. Williams, "Regularization methods for near-field acoustical holography," *J. Acoust. Soc. Am.* **110**, 1976–1988 (2001).
- ³²A. N. Tikhonov, "On stability of inverse problems," *Dokl. Akad. Nauk SSSR* **39**, 176–179 (1943).
- ³³P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems* (SIAM, Philadelphia, PA, 1998).

Analytical approach for sound attenuation in perforated dissipative silencers with inlet/outlet extensions

A. Selamet,^{a)} M. B. Xu, and I.-J. Lee

Department of Mechanical Engineering and The Center for Automotive Research, The Ohio State University, Columbus, Ohio 43210

N. T. Huff

Owens Corning, Novi, Michigan 48377

(Received 7 April 2004; revised 14 December 2004; accepted 15 January 2005)

The acoustic attenuation performance of perforated dissipative circular expansion chambers with inlet/outlet extensions is investigated. The eigenvalues and eigenfunctions of the sound field are analytically determined in the extended inlet/outlet circular ducts, upstream/downstream end annular dissipative chambers, and the central perforated dissipative expansion chamber. Utilizing the continuity conditions of velocity/pressure at the interfaces, the transmission loss is predicted by a two-dimensional analytical approach. For a specific configuration, such predictions are compared with both experiments and a three-dimensional computational solution based on the substructure boundary element technique, showing a reasonable agreement. The analytical results for the effect of the absorbent resistivity, duct porosity, and geometry on the acoustic attenuation performance are discussed in detail. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1867884]

PACS numbers: 43.50.Gf, 43.20.Mv [ANN]

Pages: 2078–2089

I. INTRODUCTION

Due to the relative simplicity of their flow path leading to low backpressure and desirable sound attenuation performance with increasing frequency, dissipative silencers have been used increasingly in gas flows, such as heating, ventilation, air conditioning, and automotive exhaust systems. To predict their acoustic behavior, both three-dimensional (3D) finite-element (FEM)^{1–3} and boundary element methods (BEM)^{4,5} have been developed for complex geometries. For silencers of axially uniform and arbitrary cross sections, various approaches^{6–11} have been adopted to reduce the computational time associated with a 3D FEM/BEM. Including the effect of mean flow and calculating the eigenvalues by FEM, Astley and Cummings⁶ obtained the modal attenuation rate in a silencer of rectangular cross section. Cummings⁷ further predicted the sound transmission in a dissipative silencer of arbitrary cross section by using a segmented Rayleigh–Ritz method. Glav obtained the eigenvalues (wave numbers) and eigenfunctions of an infinite cylindrical dissipative silencer of arbitrary cross section by using the point-matching⁸ and null-field approaches;⁹ he further expanded¹⁰ the acoustic field in a dissipative silencer by using the eigenfunctions and adjusted to the boundary conditions at the inlet/outlet by mode matching, then derived the transfer matrix for a silencer of arbitrary cross section, while providing numerical results for a silencer with elliptical cross section. Kirby¹¹ derived the eigenvalues/eigenfunctions by employing a 2D FEM, matched the expanded acoustic pressure and velocity fields in the expansion chamber of dissipative silencer to the inlet/outlet pipes by using point collocation, and obtained the transmission loss.

For a dissipative silencer of simple cross section, such as

circular, various analytical approaches have also been developed.^{12–15} Cummings and Chang¹² obtained the transmission loss of a dissipative expansion chamber including the mean flow by a mode-matching technique. Kirby¹³ investigated the acoustic attenuation of a circular dissipative silencer and obtained the transmission loss by using a simplified approach. Xu *et al.*¹⁴ developed a 2D analytical solution to examine the acoustic performance of dissipative expansion chambers by a pressure- and velocity-matching technique.¹⁶ Selamet *et al.*¹⁵ further investigated the effect of perforated ducts on the sound attenuation in dissipative silencers.

The foregoing studies^{1–15} have demonstrated that the acoustic performance of a dissipative expansion chamber can be improved by the fibrous material, though usually at relatively high frequencies. Due to the coupling of broadband modes of a simple expansion chamber and the resonant peaks of a quarter-wave resonator, the inlet/outlet extensions can enhance the transmission loss in an expansion chamber, particularly at low frequencies.^{17–21} Utilizing the mode-matching technique, Abom¹⁷ developed a general 3D analytical approach to determine the four-pole parameters including higher order mode effects for expansion chamber mufflers with inlet/outlet extensions. Selamet and Ji¹⁸ investigated the acoustic attenuation performance of circular expansion chambers with extended inlet/outlet, and employed three approaches to determine the transmission loss: a 2D axisymmetric analytical solution, a 3D BEM, and experiments. For a circular expansion chamber with extended end inlet and side outlet, they¹⁹ also investigated the acoustic attenuation and determined the transmission loss by a 1D analytical approach, a 3D BEM, and experiments. Furthermore, Selamet *et al.*²⁰ studied the acoustic behavior of a circular dual-chamber muffler by a 2D analytical approach, FEM, and experiments, and investigated the effect of rigid baffles, the

^{a)}Electronic mail: selamet.1@osu.edu

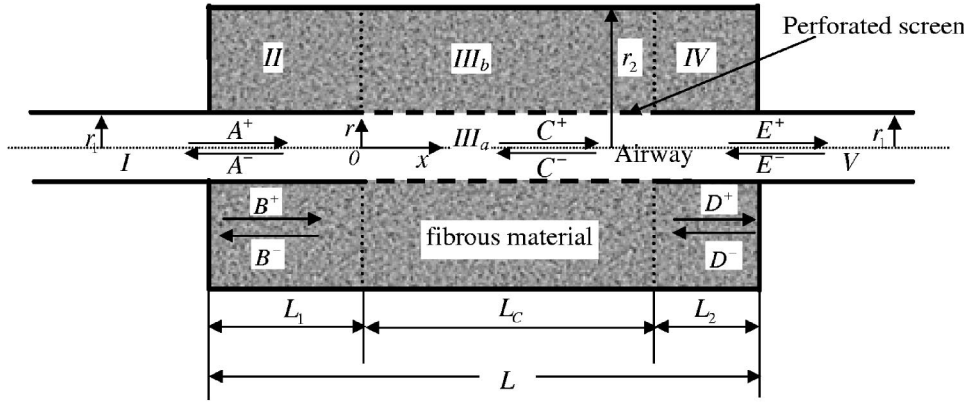


FIG. 1. Perforated dissipative circular expansion chamber with inlet/outlet extensions.

inner radius and the axial position of the baffle, and the inlet/outlet extensions. For extended-duct and perforated-duct mufflers, Torregrosa *et al.*²¹ obtained the end corrections by a 2D FEM, and validated the predictions experimentally.

Although the dissipative silencers and the expansion chambers with inlet/outlet extensions have been investigated in detail, the combination of the two concepts—the perforated dissipative silencers with inlet/outlet extensions—remains to be explored. The objective of the present study is then to: (1) investigate the acoustic attenuation of a perforated dissipative silencer with inlet/outlet extensions by a two-dimensional analytical approach; (2) compare such predictions with the experimental and numerical results for a specific configuration; and (3) examine analytically the effect of absorbent resistivity, duct porosity, and geometry on the transmission loss. Following this Introduction, Sec. II develops a two-dimensional analytical approach for the transmission loss of the perforated dissipative silencers with extended inlet/outlet. Section III discusses the effects of absorbent resistivity, duct porosity, and geometry on the transmission loss, and compares the analytical results with the numerical predictions and experiments. The study is concluded with final remarks in Sec. IV.

II. TWO-DIMENSIONAL ANALYTICAL APPROACH

Consider a cylindrical expansion chamber of radius r_2 and total length $L = L_1 + L_C + L_2$ with solid inlet (of length L_1) and outlet (of length L_2) extensions, a perforated duct of radius r_1 , length L_C , and porosity ϕ , and a sound-absorbing material between radii r_1 and r_2 , as shown in Fig. 1. The origin of the x axis is located at the end of the solid inlet duct. The absorbing material is assumed to be homogeneous and isotropic, and characterized by the complex speed of sound \tilde{c} and density $\tilde{\rho}$ (The expressions are deferred to Appendix A.) The air in the airway is characterized by the speed of sound c_0 and density ρ_0 . The extended inlet duct, the upstream end concentric annular chamber ($-L_1 \leq x \leq 0$), the central perforated expansion chamber ($0 \leq x \leq L_C$), the downstream end concentric annular chamber ($L_C \leq x \leq L_C + L_2$), and the extended outlet duct are denoted by domains I, II, III (including the central airway III_a and the concentric outer chamber with sound-absorbing material III_b), IV, and V, respectively.

A. Wave propagation in domain I

For an axisymmetric wave propagation in the extended inlet duct (domain I in Fig. 1), the solution of the Helmholtz equation can be written as

$$P_A(r, x) = \sum_{n=0}^{\infty} (A_n^+ e^{-jk_{x,A,n}x} + A_n^- e^{jk_{x,A,n}x}) \psi_{A,n}(r), \quad (1)$$

with $j = \sqrt{-1}$, P_A being the acoustic pressure in the inlet pipe; A_n^+ and A_n^- the modal amplitudes corresponding to components traveling in the positive and negative x directions, respectively; r designates the radial direction; $k_{x,A,n}$ the wave number in the x direction with subscripts x, A, n denoting axial direction, domain I, and order of the waves, respectively;

$$\psi_{A,n}(r) = J_0(k_{r,A,n}r) \quad (2)$$

is the eigenfunction, J_0 is the zeroth-order Bessel function of the first kind, and $k_{r,A,n}$ the radial wave number satisfying the rigid wall boundary condition of

$$J_0'(k_{r,A,n}r_1) = 0, \quad (3)$$

with wave numbers $k_{x,A,n}$ and $k_{r,A,n}$ being related by

$$k_{x,A,n}^2 + k_{r,A,n}^2 = k_0^2, \quad (4)$$

where $k_0 = \omega/c_0 = 2\pi f/c_0$ is the wave number in air, ω being the angular velocity, and f the frequency. In view of the linearized momentum equation, the particle velocity in the axial direction may then be written as

$$u_{x,A}(r, x) = \frac{1}{\rho_0 \omega} \sum_{n=0}^{\infty} k_{x,A,n} (A_n^+ e^{-jk_{x,A,n}x} - A_n^- e^{jk_{x,A,n}x}) \times \psi_{A,n}(r). \quad (5)$$

B. Wave propagation in domain II

The pressure for a two-dimensional axisymmetric wave propagation in the upstream annular duct (Fig. 1: domain II of length L_1 , inner radius r_1 , outer radius r_2 , and filled with sound-absorbing material) is expressed as

$$P_B(r, x) = \sum_{n=0}^{\infty} (B_n^+ e^{-jk_{x,B,n}x} + B_n^- e^{jk_{x,B,n}x}) \psi_{B,n}(r), \quad (6)$$

with P_B being the acoustic pressure; B_n^+ and B_n^- the modal amplitudes corresponding to components traveling in the positive and negative x directions, respectively; $k_{x,B,n}$ the axial wave number with subscript B denoting domain II. The eigenfunction for such a circular, concentric annular duct is

$$\psi_{B,n}(r) = J_0(k_{r,B,n}r) - [J_1(k_{r,B,n}r_2)/Y_1(k_{r,B,n}r_2)]Y_0(k_{r,B,n}r), \quad (7)$$

with J_1 being the first-order Bessel function of the first kind; Y_0 and Y_1 being Bessel functions of the second kind of zeroth- and first order; $k_{r,B,n}$ the radial wave number satisfying

$$J_1(k_{r,B,n}r_1) - [J_1(k_{r,B,n}r_2)/Y_1(k_{r,B,n}r_2)]Y_1(k_{r,B,n}r_1) = 0. \quad (8)$$

The axial wave number of the n mode $k_{x,B,n}$ is related to $k_{r,B,n}$ by

$$k_{r,B,n}^2 + k_{x,B,n}^2 = \tilde{k}^2, \quad (9)$$

with $\tilde{k} = 2\pi f/\tilde{c}$ being the wave number in the fibrous material. From the linearized momentum equation, Eq. (6) yields the particle velocity in the axial direction as

$$u_{x,B}(r,x) = \frac{1}{\tilde{\rho}\omega} \sum_{n=0}^{\infty} k_{x,B,n} (B_n^+ e^{-jk_{x,B,n}x} - B_n^- e^{jk_{x,B,n}x}) \times \psi_{B,n}(r). \quad (10)$$

C. Wave propagation in domains III_a and III_b

Consider the central perforated dissipative expansion chamber of radius r_2 (domain III in Fig. 1) with a perforated screen of radius r_1 and porosity ϕ that separates the central airway (domain III_a), and the sound-absorbing material in the outer chamber between radii r_1 and r_2 (domain III_b). The pressure may then be expressed as

$$P_C(r,x) = \sum_{n=0}^{\infty} (C_n^+ e^{-jk_{x,C,n}x} + C_n^- e^{jk_{x,C,n}x}) \psi_{C,n,P}(r), \quad (11)$$

with

$$P_C(r,x) = \begin{cases} P_{III_a}(r,x), & 0 \leq r \leq r_1 \\ P_{III_b}(r,x), & r_1 \leq r \leq r_2 \end{cases}, \quad (12a)$$

and

$$\psi_{C,n,P}(r) = \begin{cases} \psi_{III_a,n,P}(r), & 0 \leq r \leq r_1 \\ \psi_{III_b,n,P}(r), & r_1 \leq r \leq r_2 \end{cases}, \quad (12b)$$

where subscripts C , III_a, III_b denote domains III, III_a, and III_b, respectively; C_n^+ and C_n^- the modal amplitudes; $k_{x,C,n}$ the axial wave number for both the central airway and absorbent material. The transverse modal eigenfunction for the pressure is given by

$$\psi_{C,n,P}(r) = \begin{cases} \psi_{III_a,n,P}(r) = C_1 J_0(k_{r,C,n}r) + C_2 Y_0(k_{r,C,n}r), & 0 \leq r \leq r_1 \\ \psi_{III_b,n,P}(r) = C_3 J_0(\tilde{k}_{r,C,n}r) + C_4 Y_0(\tilde{k}_{r,C,n}r), & r_1 \leq r \leq r_2 \end{cases}, \quad (13)$$

where $k_{r,C,n}$ and $\tilde{k}_{r,C,n}$ are the radial wave numbers for the air and fibrous material, respectively, which are related to $k_{x,C,n}$ by

$$k_{r,C,n}^2 + k_{x,C,n}^2 = k_0^2, \quad (14a)$$

$$\tilde{k}_{r,C,n}^2 + k_{x,C,n}^2 = \tilde{k}^2. \quad (14b)$$

The wave numbers can be obtained from the characteristic equation (derived from the boundary conditions of acoustic pressure and radial particle velocity at $r=0$, r_1 , and r_2) as¹⁵

$$\frac{\rho_0 \tilde{k}_{r,C,n}}{\tilde{\rho} k_{r,C,n}} \left[\frac{J_0(k_{r,C,n}r_1)}{J_1(k_{r,C,n}r_1)} + \frac{j\tilde{\xi}_p k_{r,C,n}}{k_0} \right] = \frac{J_0(\tilde{k}_{r,C,n}r_1)Y_1(\tilde{k}_{r,C,n}r_2) - Y_0(\tilde{k}_{r,C,n}r_1)J_1(\tilde{k}_{r,C,n}r_2)}{J_1(\tilde{k}_{r,C,n}r_1)Y_1(\tilde{k}_{r,C,n}r_2) - Y_1(\tilde{k}_{r,C,n}r_1)J_1(\tilde{k}_{r,C,n}r_2)}, \quad (15)$$

with $\tilde{\xi}_p$ (given in Appendix B) being the nondimensionalized perforate acoustic impedance, which relates the acoustic pressure in the inner duct and outer chamber through the interface. The coefficients $C_1 - C_4$ can also be expressed as

$$C_1 = 1, \quad (16a)$$

$$C_2 = 0, \quad (16b)$$

$$C_3 = \frac{k_{r,C,n}\tilde{\rho}}{\tilde{k}_{r,C,n}\rho_0} \times \frac{J_1(k_{r,C,n}r_1)Y_1(\tilde{k}_{r,C,n}r_2)}{J_1(\tilde{k}_{r,C,n}r_1)Y_1(\tilde{k}_{r,C,n}r_2) - J_1(\tilde{k}_{r,C,n}r_2)Y_1(\tilde{k}_{r,C,n}r_1)}, \quad (16c)$$

$$C_4 = -\frac{J_1(\tilde{k}_{r,C,n}r_2)}{Y_1(\tilde{k}_{r,C,n}r_2)} C_3 = -\frac{k_{r,C,n}\tilde{\rho}}{\tilde{k}_{r,C,n}\rho_0} \times \frac{J_1(k_{r,C,n}r_1)J_1(\tilde{k}_{r,C,n}r_2)}{J_1(\tilde{k}_{r,C,n}r_1)Y_1(\tilde{k}_{r,C,n}r_2) - J_1(\tilde{k}_{r,C,n}r_2)Y_1(\tilde{k}_{r,C,n}r_1)}. \quad (16d)$$

With the solution of the wave numbers from Eq. (15) and the coefficients $C_1 - C_4$ from Eq. (16), Eq. (13) yields the expression for the transverse modal eigenfunction of the pressure as

$$\psi_{C,n,p}(r) = \begin{cases} J_0(k_{r,C,n}r), & 0 \leq r \leq r_1 \\ C_5 \left[J_0(\tilde{k}_{r,C,n}r) - \frac{J_1(\tilde{k}_{r,C,n}r_2)}{Y_1(\tilde{k}_{r,C,n}r_2)} Y_0(\tilde{k}_{r,C,n}r) \right], & r_1 \leq r \leq r_2 \end{cases}, \quad (17a)$$

with

$$C_5 = \frac{Y_1(\tilde{k}_{r,C,n}r_2) \left[J_0(k_{r,C,n}r_1) + \frac{j\tilde{\xi}_p k_{r,C,n}}{k_0} J_1(k_{r,C,n}r_1) \right]}{J_0(\tilde{k}_{r,C,n}r_1) Y_1(\tilde{k}_{r,C,n}r_2) - J_1(\tilde{k}_{r,C,n}r_2) Y_0(\tilde{k}_{r,C,n}r_1)}. \quad (17b)$$

From the linearized momentum equation

$$u_{x,C}(r,x) = \begin{cases} \frac{j}{\rho_0 \omega} \frac{\partial P_{III_a}}{\partial x}, & 0 \leq r \leq r_1 \\ \frac{j}{\tilde{\rho} \omega} \frac{\partial P_{III_b}}{\partial x}, & r_1 \leq r \leq r_2 \end{cases}, \quad (18)$$

Eq. (17) yields the particle velocity in the axial direction as

$$u_{x,C}(r,x) = \frac{1}{\rho_0 \omega} \sum_{n=0}^{\infty} k_{x,C,n} (C_n^+ e^{-jk_{x,C,n}x} - C_n^- e^{jk_{x,C,n}x}) \times \psi_{C,n,u_x}(r), \quad (19)$$

with the transverse modal eigenfunction for the velocity being expressed as

$$\psi_{C,n,u_x}(r) = \begin{cases} J_0(k_{r,C,n}r), & 0 \leq r \leq r_1 \\ C_5 \frac{\rho_0}{\tilde{\rho}} \left[J_0(\tilde{k}_{r,C,n}r) - \frac{J_1(\tilde{k}_{r,C,n}r_2)}{Y_1(\tilde{k}_{r,C,n}r_2)} Y_0(\tilde{k}_{r,C,n}r) \right], & r_1 \leq r \leq r_2 \end{cases}. \quad (20)$$

D. Wave propagation in domains IV and V

The acoustic pressure and axial particle velocity in domains IV and V are similar to those in domains I and II, and can be readily expressed as

$$P_D(r,x) = \sum_{n=0}^{\infty} (D_n^+ e^{-jk_{x,D,n}(x-L_C)} + D_n^- e^{jk_{x,D,n}(x-L_C)}) \psi_{D,n}(r), \quad (21a)$$

$$P_E(r,x) = \sum_{n=0}^{\infty} (E_n^+ e^{-jk_{x,E,n}(x-L_C)} + E_n^- e^{jk_{x,E,n}(x-L_C)}) \psi_{E,n}(r), \quad (21b)$$

$$u_{x,D}(r,x) = \frac{1}{\tilde{\rho} \omega} \sum_{n=0}^{\infty} k_{x,D,n} (D_n^+ e^{-jk_{x,D,n}(x-L_C)} - D_n^- e^{jk_{x,D,n}(x-L_C)}) \psi_{D,n}(r), \quad (21c)$$

and

$$u_{x,E}(r,x) = \frac{1}{\rho_0 \omega} \sum_{n=0}^{\infty} k_{x,E,n} (E_n^+ e^{-jk_{x,E,n}(x-L_C)} - E_n^- e^{jk_{x,E,n}(x-L_C)}) \psi_{E,n}(r), \quad (21d)$$

where subscripts D and E denote domains IV and V, respectively; D_n^+ , D_n^- , E_n^+ , and E_n^- the amplitudes; with both the

eigenfunctions $\psi_{D,n}(r)$, $\psi_{E,n}(r)$ and wave numbers $k_{x,D,n}$, $k_{x,E,n}$ being similar to those of domains II and I.

E. Transmission loss prediction

With the pressure and particle velocity given by Eqs. (1), (5), (6), (10), (17), (19), and (21), transmission loss can then be obtained by solving the unknown coefficients A_n , B_n , C_n , D_n , and E_n using the boundary conditions at $x = -L_1$, 0 , L_C , and $L_C + L_2$. In view of Eqs. (10) and (21), the rigid boundary condition at the left ($x = -L_1$) and right ($x = L_C + L_2$) endplate of the chamber

$$u_{x,B} = 0, \quad \text{for } r_1 \leq r \leq r_2, \quad x = -L_1, \quad (22a)$$

and

$$u_{x,D} = 0, \quad \text{for } r_1 \leq r \leq r_2, \quad x = L_C + L_2, \quad (22b)$$

yields

$$B_n^+ = B_n^- e^{-2jk_{x,B,n}L_1}, \quad (23a)$$

$$D_n^- = D_n^+ e^{-2jk_{x,D,n}L_2}. \quad (23b)$$

At the interfaces of expansion ($x=0$) and contraction ($x=L_C$), the continuities of the acoustic pressure and axial velocity give

$$P_C = P_A, \quad \text{for } 0 \leq r \leq r_1, \quad x=0, \quad (24a)$$

$$P_C = P_B, \quad \text{for } r_1 \leq r \leq r_2, \quad x=0, \quad (24b)$$

$$u_{x,C} = \begin{cases} u_{x,A}, & \text{for } 0 \leq r \leq r_1, x=0 \\ u_{x,B}, & \text{for } r_1 \leq r \leq r_2, x=0 \end{cases} \quad (24c)$$

and

$$P_C = P_E, \quad \text{for } 0 \leq r \leq r_1, x=L_C, \quad (25a)$$

$$P_C = P_D, \quad \text{for } r_1 \leq r \leq r_2, x=L_C, \quad (25b)$$

$$u_{x,C} = \begin{cases} u_{x,E}, & \text{for } 0 \leq r \leq r_1, x=L_C \\ u_{x,D}, & \text{for } r_1 \leq r \leq r_2, x=L_C \end{cases} \quad (25c)$$

In view of the expressions of the pressure and velocity as infinite series of unknown amplitudes in Eqs. (1), (5), (6), (10), (17), (20), and (21), Eqs. (24) and (25) yield

$$\begin{aligned} & \sum_{n=0}^{\infty} (C_n^+ + C_n^-) \psi_{C,n,P}(r) \\ &= \sum_{n=0}^{\infty} (A_n^+ + A_n^-) \psi_{A,n}(r), \quad \text{for } 0 \leq r \leq r_1, \end{aligned} \quad (26a)$$

$$\begin{aligned} & \sum_{n=0}^{\infty} (C_n^+ + C_n^-) \psi_{C,n,P}(r) \\ &= \sum_{n=0}^{\infty} (B_n^+ + B_n^-) \psi_{B,n}(r), \quad \text{for } r_1 \leq r \leq r_2, \end{aligned} \quad (26b)$$

$$\begin{aligned} & \sum_{n=0}^{\infty} k_{x,C,n} (C_n^+ - C_n^-) \psi_{C,n,u_x}(r) \\ &= \begin{cases} \sum_{n=0}^{\infty} k_{x,A,n} (A_n^+ - A_n^-) \psi_{A,n}(r), & \text{for } 0 \leq r \leq r_1 \\ \frac{\rho_0}{\bar{\rho}} \sum_{n=0}^{\infty} k_{x,B,n} (B_n^+ - B_n^-) \psi_{B,n}(r), & \text{for } r_1 \leq r \leq r_2 \end{cases}, \end{aligned} \quad (26c)$$

and

$$\begin{aligned} & \sum_{n=0}^{\infty} (C_n^+ e^{-jk_{x,C,n}L_C} + C_n^- e^{jk_{x,C,n}L_C}) \psi_{C,n,P}(r) \\ &= \sum_{n=0}^{\infty} (E_n^+ + E_n^-) \psi_{E,n}(r), \quad \text{for } 0 \leq r \leq r_1, \end{aligned} \quad (27a)$$

$$\begin{aligned} & \sum_{n=0}^{\infty} (C_n^+ e^{-jk_{x,C,n}L_C} + C_n^- e^{jk_{x,C,n}L_C}) \psi_{C,n,P}(r) \\ &= \sum_{n=0}^{\infty} (D_n^+ + D_n^-) \psi_{D,n}(r), \quad \text{for } r_1 \leq r \leq r_2, \end{aligned} \quad (27b)$$

$$\begin{aligned} & \sum_{n=0}^{\infty} k_{x,C,n} (C_n^+ e^{-jk_{x,C,n}L_C} - C_n^- e^{jk_{x,C,n}L_C}) \psi_{C,n,u_x}(r) \\ &= \begin{cases} \sum_{n=0}^{\infty} k_{x,E,n} (E_n^+ - E_n^-) \psi_{E,n}(r), & \text{for } 0 \leq r \leq r_1 \\ \frac{\rho_0}{\bar{\rho}} \sum_{n=0}^{\infty} k_{x,D,n} (D_n^+ - D_n^-) \psi_{D,n}(r), & \text{for } r_1 \leq r \leq r_2 \end{cases}. \end{aligned} \quad (27c)$$

In order to solve Eqs. (26) and (27), the infinite series of unknown amplitudes needs to be truncated to a suitable number. In solving a similar problem, the acoustic field across the discontinuities was matched in Refs. 14 and 15 by imposing the continuities of the acoustic pressure/velocity integral over discrete zones at the expansion/contraction. For dissipative silencers with inlet/outlet extensions, another analytical approach¹⁸ proves to be more suitable to match the acoustic field across the interfaces of expansion/contraction. For the continuities of pressure at interfaces of $x=0$ and L_C , multiplying both sides of Eqs. (26a) and (27a) by $\psi_{A,s}dS$ or $\psi_{E,s}dS$ gives

$$\begin{aligned} & \sum_{n=0}^{\infty} (C_n^+ + C_n^-) \langle \psi_{C,n,P} \psi_{A,s} \rangle_{0 \leq r \leq r_1} \\ &= \sum_{n=0}^{\infty} (A_n^+ + A_n^-) \langle \psi_{A,n} \psi_{A,s} \rangle_{0 \leq r \leq r_1}, \quad \text{for } s=0,1,\dots,\infty, \end{aligned} \quad (28)$$

$$\begin{aligned} & \sum_{n=0}^{\infty} (C_n^+ e^{-jk_{x,C,n}L_C} + C_n^- e^{jk_{x,C,n}L_C}) \langle \psi_{C,n,P} \psi_{E,s} \rangle_{0 \leq r \leq r_1} \\ &= \sum_{n=0}^{\infty} (E_n^+ + E_n^-) \langle \psi_{E,n} \psi_{E,s} \rangle_{0 \leq r \leq r_1}, \quad \text{for } s=0,1,\dots,\infty, \end{aligned} \quad (29)$$

with $\langle \rangle$ being the integration over the surface. Similarly, by multiplying $\psi_{B,s}dS$ or $\psi_{D,s}dS$, Eqs. (26b) and (27b) give

$$\begin{aligned} & \sum_{n=0}^{\infty} (C_n^+ + C_n^-) \langle \psi_{C,n,P} \psi_{B,s} \rangle_{r_1 \leq r \leq r_2} \\ &= \sum_{n=0}^{\infty} (B_n^+ + B_n^-) \langle \psi_{B,n} \psi_{B,s} \rangle_{r_1 \leq r \leq r_2}, \quad \text{for } s=0,1,\dots,\infty, \end{aligned} \quad (30)$$

$$\begin{aligned} & \sum_{n=0}^{\infty} (C_n^+ e^{-jk_{x,C,n}L_C} + C_n^- e^{jk_{x,C,n}L_C}) \langle \psi_{C,n,P} \psi_{D,s} \rangle_{r_1 \leq r \leq r_2} \\ &= \sum_{n=0}^{\infty} (D_n^+ + D_n^-) \langle \psi_{D,n} \psi_{D,s} \rangle_{r_1 \leq r \leq r_2}, \quad \text{for } s=0,1,\dots,\infty. \end{aligned} \quad (31)$$

By multiplying both sides of Eqs. (26c) and (27c) by $\psi_{C,s,u_x}dS$, the continuities of axial velocity at interfaces of $x=0$ and L_C provide

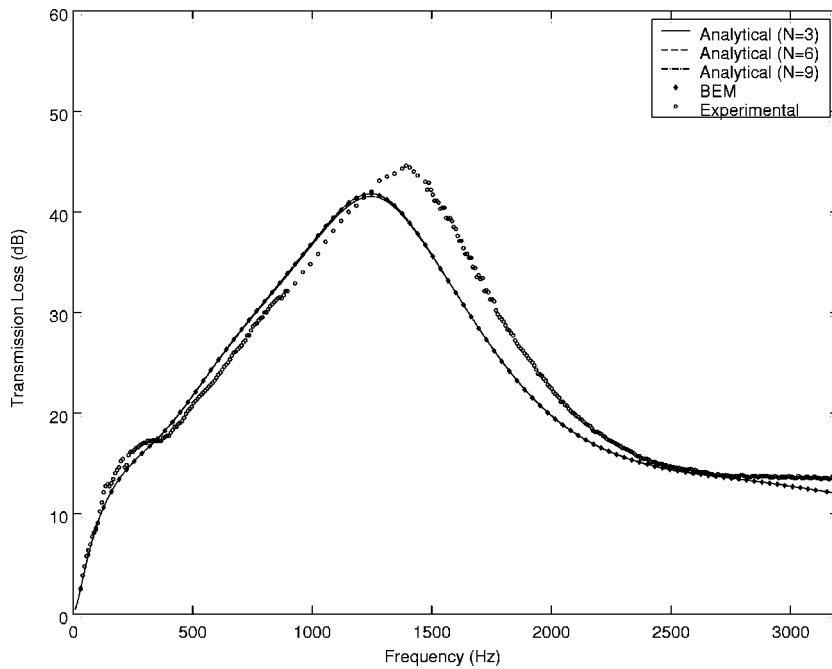


FIG. 2. Transmission loss of a perforated dissipative silencer with inlet/outlet extensions ($r_1=2.45$ cm, $r_2=8.22$ cm, $L=L_1+L_C+L_2=25.72$ cm, $L_1=L_2=2.4$ cm, $R=4896$ Rayls/m, $\phi=8\%$, $d_h=0.249$ cm, and $t_w=0.09$ cm).

$$\begin{aligned} & \sum_{n=0}^{\infty} k_{x,C,n}(C_n^+ - C_n^-) \langle \psi_{C,n,u_x} \psi_{C,s,u_x} \rangle_{0 \leq r \leq r_2} \\ &= \sum_{n=0}^{\infty} k_{x,A,n}(A_n^+ - A_n^-) \langle \psi_{A,n} \psi_{C,s,u_x} \rangle_{0 \leq r \leq r_1} \\ &+ \frac{\rho_0}{\bar{\rho}} \sum_{n=0}^{\infty} k_{x,B,n}(B_n^+ - B_n^-) \langle \psi_{B,n} \psi_{C,s,u_x} \rangle_{r_1 \leq r \leq r_2}, \\ & \text{for } s=0,1,\dots,\infty, \end{aligned} \quad (32)$$

$$\begin{aligned} & \sum_{n=0}^{\infty} k_{x,C,n}(C_n^+ e^{-jk_{x,C,n}L_C} - C_n^- e^{jk_{x,C,n}L_C}) \\ & \times \langle \psi_{C,n,u_x} \psi_{C,s,u_x} \rangle_{0 \leq r \leq r_2} \\ &= \sum_{n=0}^{\infty} k_{x,E,n}(E_n^+ - E_n^-) \langle \psi_{E,n} \psi_{C,s,u_x} \rangle_{0 \leq r \leq r_1} \\ &+ \frac{\rho_0}{\bar{\rho}} \sum_{n=0}^{\infty} k_{x,D,n}(D_n^+ - D_n^-) \langle \psi_{D,n} \psi_{C,s,u_x} \rangle_{r_1 \leq r \leq r_2}, \\ & \text{for } s=0,1,\dots,\infty. \end{aligned} \quad (33)$$

The expressions of the integrals $\langle \rangle$ in Eqs. (28)–(33) are provided in Appendix C.

To determine the transmission loss of the perforated dissipative silencers with inlet/outlet extensions: (1) the incoming wave is assumed to be planar and its magnitude A_0^+ is chosen to be unity for convenience, and (2) an anechoic termination is imposed at the exit by setting $E_n^- = 0$. Thus, in view of Eq. (23), Eqs. (28)–(33) give a large (theoretically infinite) number of relations $6(s+1)$ for a large number of unknowns $6(n+1)$. The unknowns are the magnitudes of incident and reflected waves in domains I–V (A_n^- , B_n^- , C_n^+ , C_n^- , D_n^+ , and E_n^+). Since higher modes have a diminishing effect on the solution, s and n can be truncated to N , resulting

in $6(N+1)$ equations with $6(N+1)$ unknowns. The value of N needed for a converged solution depends on the dimensions of the chamber, and the frequency range of interest. Once Eqs. (28)–(33) are solved, the transmission loss may be determined as

$$\text{TL} = -20 \log_{10} |E_0^+|, \quad (34)$$

under the assumption that all transmitted waves in the outlet pipe are nonpropagating modes except the first mode.

III. RESULTS AND DISCUSSION

For the perforated, dissipative circular expansion chamber with inlet/outlet extensions studied here (recall Fig. 1), the following base parameters are chosen: $r_1=2.45$ cm, $r_2=8.22$ cm, $L=L_1+L_C+L_2=25.72$ cm, $R=4896$ Rayls/m (for a filling density of 100 g/l), porosity of perforated duct $\phi=8\%$, thickness of the perforated duct $t_w=0.09$ cm, and diameter of the holes $d_h=0.249$ cm. With the exception of solid inlet/outlet extensions, this choice retains consistency with the configuration of Ref. 15.

Figure 2 presents the analytical results for the transmission loss of a perforated dissipative circular expansion chamber with extensions $L_1=L_2=2.4$ cm. To examine the accuracy of the analytical approach, the results from experiments (utilizing the two-microphone technique) and computations (BEM⁵) are also provided. For the current geometry and frequency range of interest, the analytical results are sufficiently accurate for $N \geq 9$; hence, $N=9$ is used in the remainder of the study. The analytical results agree well with those of BEM, and show a reasonable comparison with the measurements. The transmission loss of this silencer has a peak value at the resonance frequency around 1200 Hz, and exceeds 30 dB at a relatively broad frequency range ($770 \text{ Hz} \leq f \leq 1650 \text{ Hz}$.) The difference between predictions and measurements is attributed to the approximations in-

volved in expressions for absorbing material properties and perforation impedance, and filling/texturization method used in the experiment.

Analytical TL predictions for perforated dissipative expansion chambers with different lengths of inlet/outlet extensions (L_1, L_2) are presented in Fig. 3, while retaining the same total length (L). Figure 3(a) shows results for silencers with $L_2=0$ and varying L_1 ; Fig. 3(b) with $L_1=L_2$ and varying lengths; Fig. 3(c) with $L_1 \neq L_2$ and varying lengths. The no-extension ($L_1=L_2=0$) base case is also included in these figures for comparison. Figures 3(a)–(c) reveal that: (1) the extensions do not influence the low-frequency behavior, as could partially be inferred from earlier empty silencer results;¹⁸ (2) increasing extension lengths reduces TL at high frequencies; and (3) relatively small extensions improve TL in midfrequency range: Note, for example, the silencer of Fig. 3(b) with $L_1=L_2=3.5$ cm with improved attenuation (40 dB at the frequency range of $980 \text{ Hz} \leq f \leq 1450 \text{ Hz}$). Further increase in extension lengths may suppress the attenuation relative to the no-extension case: Note, for example, the $L_1=L_2=7.5$ cm configuration in Fig. 3(b). The results in Figures 3(a)–(c) for filled silencers may be contrasted to their counterparts in Figs. 4(a)–(c) with the filling and perforations removed, while retaining the same geometry.

Using the same pairs of inlet/outlet extensions in Fig. 3, the attenuation performance in the absence of perforated screen is presented in Fig. 5. The effect of extensions at low- and high-frequency regions are qualitatively similar to Fig. 3, while the reduction in TL is smaller in the absence of perforations. Attenuation in midrange frequencies shows some qualitative similarity to that of Fig. 3, for example, in terms of some augmentation with shorter extensions and deterioration at longer extensions [see, for example, Fig. 5(b)]. The same midrange also exhibits the differences such as: (1) the longer extensions in the absence of perforations do not hinder the TL as much; and (2) the resonances tend to move to different frequencies and length combinations.

Analytical transmission loss for perforated dissipative silencers with different flow resistivities ($R=1000, 4896$, and $17\,378$ Rayls/m) are presented in Fig. 6, as contrasted to the limiting case of an empty chamber. The relative behavior is frequency dependent, with higher flow resistivity reducing TL at low frequencies and improving at high frequencies. The performance of the base model (with $R=4896$ Rayls/m) is better than the other three silencers at frequencies below 1400 Hz.

The effect of perforation porosity ($\phi=2\%$, 8% , and 50%) is shown in Fig. 7, in comparison with the limiting case of removed perforations. High porosity improves the attenuation at relatively high frequencies, whereas the lower porosities ($\phi=2\%$ and 8%) introduce resonances at low to midfrequency range, thereby improving TL. The behavior of the silencer with $\phi=2\%$ resembles that of a Helmholtz resonator, while the one with $\phi=50\%$ exhibits a behavior similar to that of no perforations, as expected.

IV. CONCLUDING REMARKS

The acoustic attenuation performance of a perforated single-pass, concentric cylindrical dissipative expansion

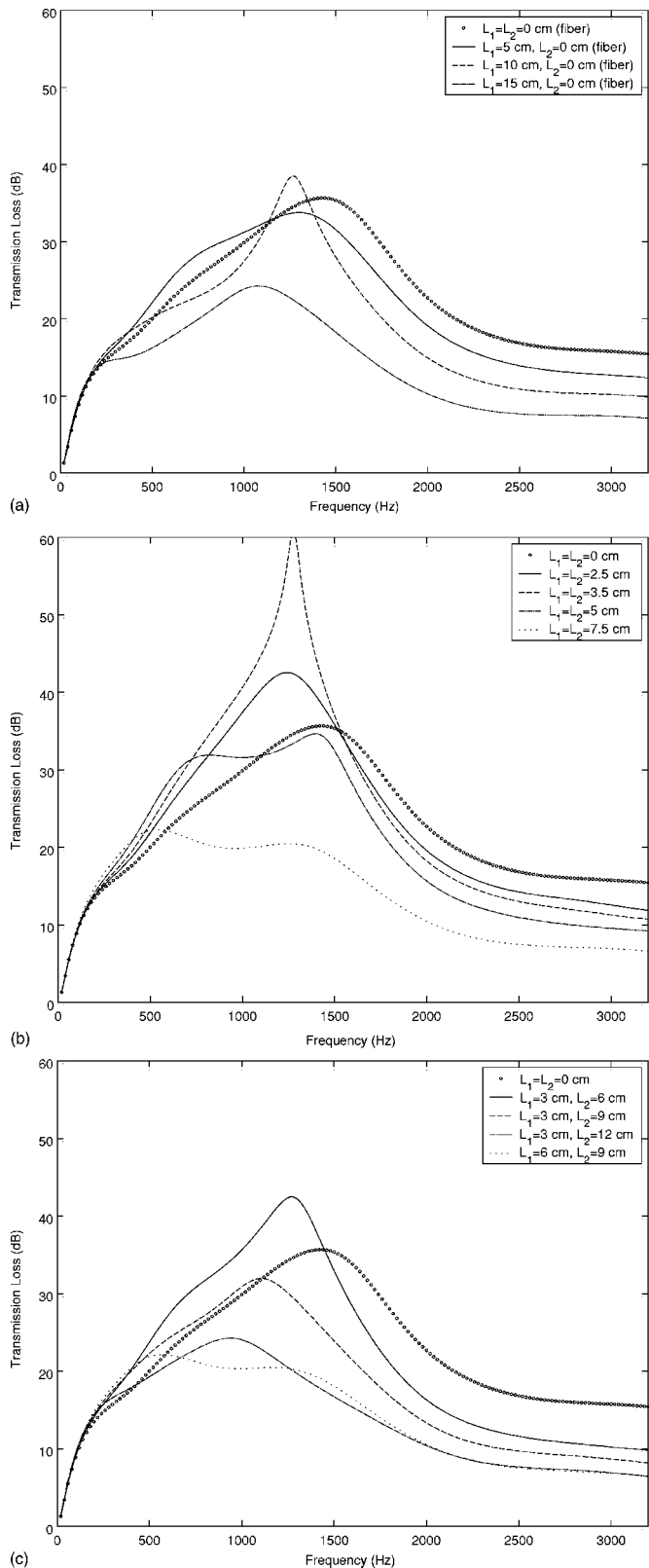


FIG. 3. Transmission loss of perforated dissipative silencers with varying lengths of inlet/outlet extensions ($r_1=2.45$ cm, $r_2=8.22$ cm, $L=25.72$ cm, $R=4896$ Rayls/m, $\phi=8\%$, $d_h=0.249$ cm, and $t_w=0.09$ cm). (a) $L_2=0$ and varying L_1 ; (b) $L_1=L_2$ and varying lengths; (c) $L_1 \neq L_2$ and varying lengths.

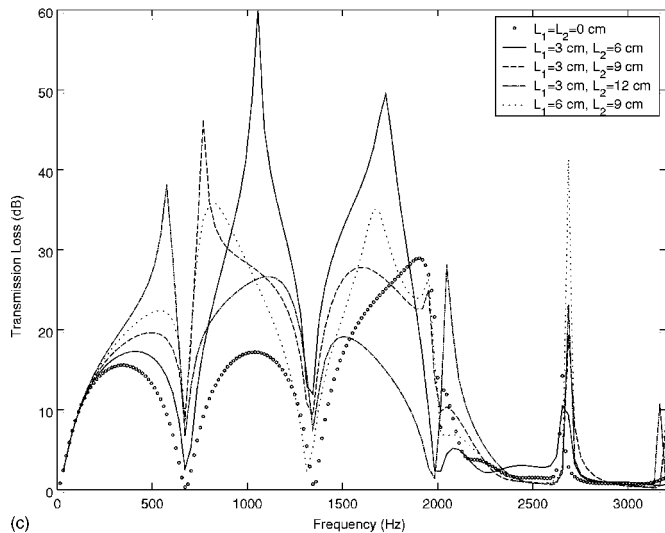
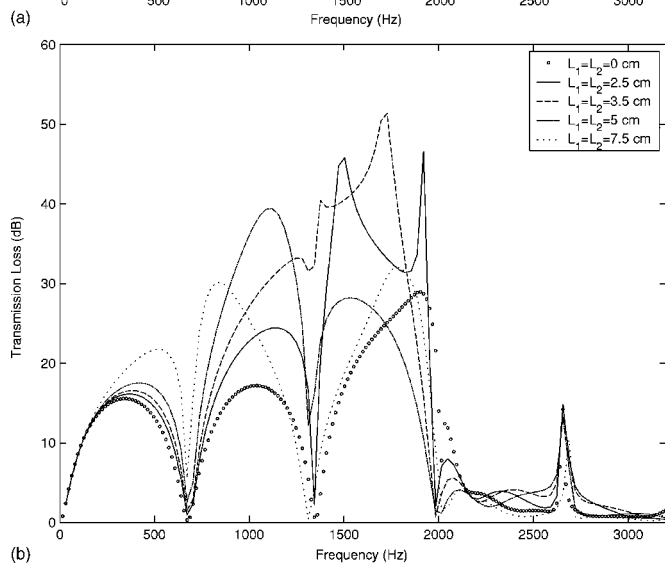
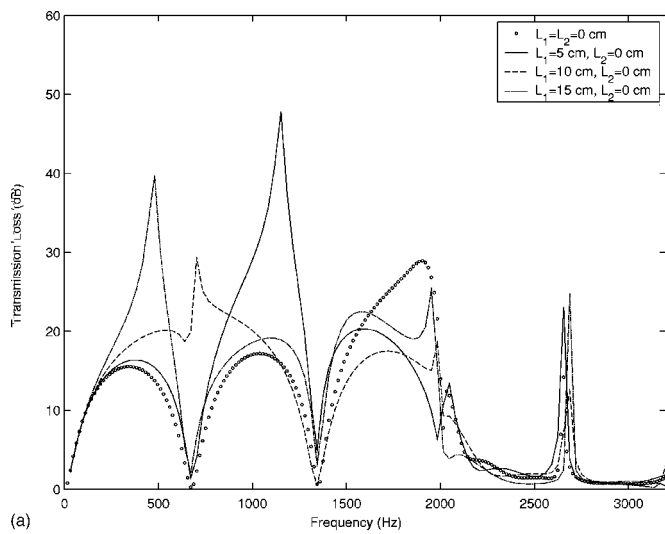


FIG. 4. Transmission loss of expansion chamber type silencers with varying lengths of inlet/outlet extensions ($r_1=2.45$ cm, $r_2=8.22$ cm, $L=25.72$ cm). (a) $L_2=0$ and varying L_1 ; (b) $L_1=L_2$ and varying lengths; (c) $L_1 \neq L_2$ and varying lengths.

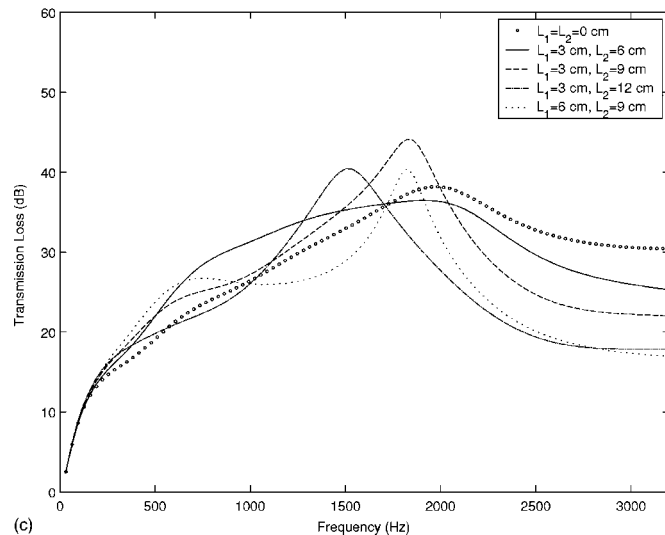
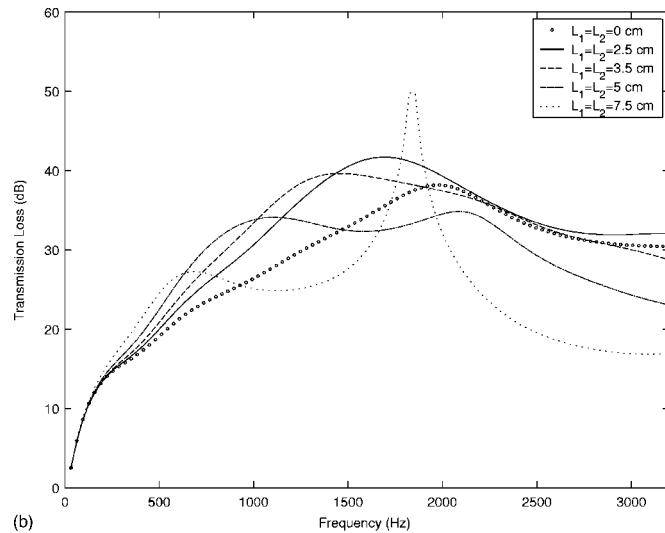
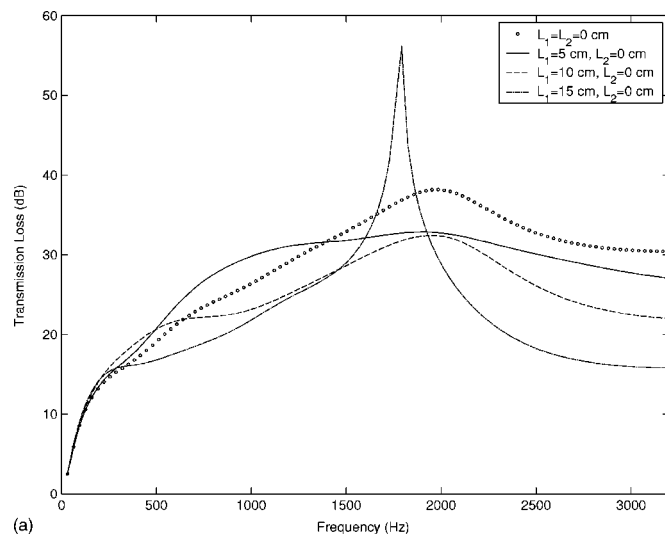


FIG. 5. Effect of inlet/outlet extensions length on the transmission loss of dissipative silencers without perforated screen ($r_1=2.45$ cm, $r_2=8.22$ cm, $L=25.72$ cm, $R=4896$ Rayls/m). (a) $L_2=0$ and varying L_1 ; (b) $L_1=L_2$ and varying lengths; (c) $L_1 \neq L_2$ and varying lengths.

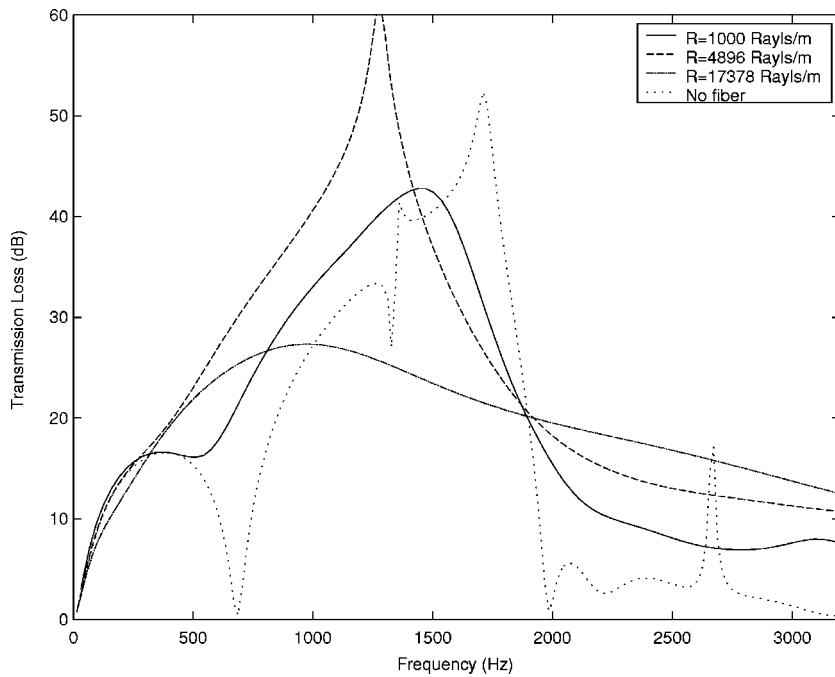


FIG. 6. Transmission loss of perforated dissipative extended silencers with different flow resistivities ($r_1 = 2.45$ cm, $r_2 = 8.22$ cm, $L = 25.72$ cm, $L_1 = L_2 = 3.5$ cm, $\phi = 8\%$, $d_h = 0.249$ cm, and $t_w = 0.09$ cm).

chamber with inlet/outlet extensions has been investigated by using a two-dimensional, closed-form analytical approach. With the expressions of the acoustic eigenvalues and eigenfunctions in various domains of the silencer, the transmission loss of the silencer is obtained analytically by utilizing the continuities of the acoustic pressure and particle velocity. The expected agreement between the analytical and BEM results has been illustrated for a specific configuration. These results, in turn, compare reasonably well with the measured TL. The acoustic attenuation of dissipative silencers with inlet/outlet extensions is also compared with their counterparts with the filling or perforations removed. The effects of length extensions, absorbent resistivity, and perforation porosity on the acoustic attenuation of the silencers have been illustrated.

APPENDIX A: ACOUSTICAL PROPERTIES OF FIBROUS MATERIAL

The acoustic properties of the fibrous material are characterized by the complex-valued characteristic impedance \tilde{Z} $= \tilde{\rho}\tilde{c}$ and wave number $\tilde{k} = 2\pi f/\tilde{c}$, expressed as

$$\frac{\tilde{Z}}{Z_0} = [1 + 0.855(f/R)^{-0.754}] + j[-0.0765(f/R)^{-0.732}], \quad (\text{A1})$$

$$\frac{\tilde{k}}{k} = [1 + 0.1472(f/R)^{-0.577}] + j[-0.1734(f/R)^{-0.595}], \quad (\text{A2})$$

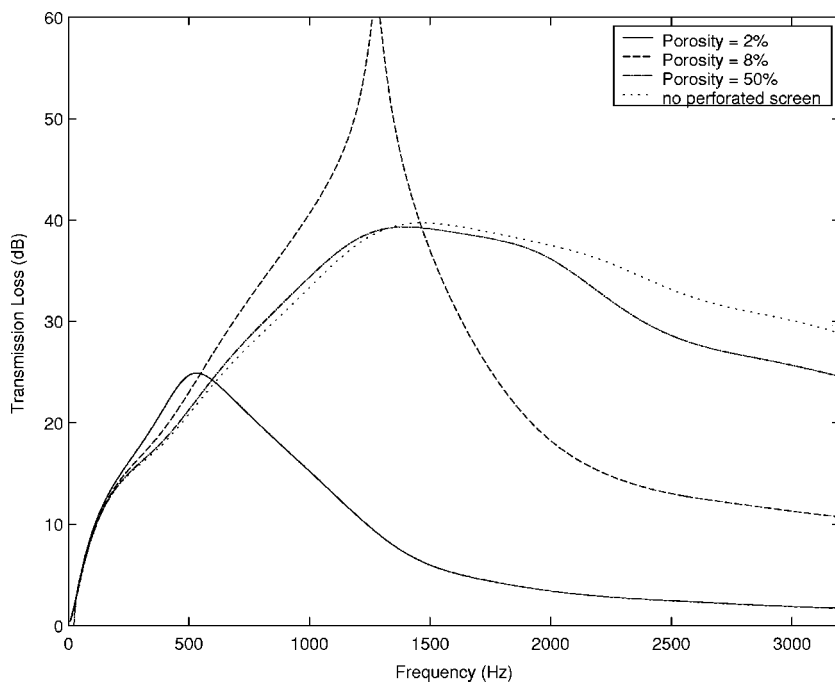


FIG. 7. Transmission loss of perforated dissipative extended silencers with different perforation porosities ($r_1 = 2.45$ cm, $r_2 = 8.22$ cm, $L = 25.72$ cm, $L_1 = L_2 = 3.5$ cm, $R = 4896$ Rayls/m, $d_h = 0.249$ cm, and $t_w = 0.09$ cm).

with $Z_0 = \rho_0 c_0$ being the characteristic impedance of the air, and R [mks Rayls/m] the flow resistivity.

$$\tilde{\zeta}_p = [0.006 + jk_0(t_w + 0.75d_h)]/\phi, \quad (\text{B1})$$

APPENDIX B: ACOUSTIC IMPEDANCE OF PERFORATES

In Eq. (15), the nondimensionalized perforate acoustic impedance $\tilde{\zeta}_p$ relates the acoustic pressure in the inner duct and outer chamber through the interface. An empirical expression for the acoustic impedance of perforate holes facing air is presented as²²⁻²⁴

with ϕ being the porosity. In the presence of absorbent (for perforations facing absorbing material), Eq. (B1) has been modified as

$$\tilde{\zeta}_p = \left[0.006 + jk_0 \left\{ t_w + 0.375d_h \left(1 + \frac{\tilde{Z}}{Z_0} \frac{\tilde{k}}{k_0} \right) \right\} \right] / \phi. \quad (\text{B2})$$

APPENDIX C: INTEGRALS IN EQS. (28)–(33)

From the integral relations for Bessel functions of any kind, $B_0(k_s r)$ and $B_0(k_n r)$

$$\int r B_0(k_s r) B_0(k_n r) dr = \begin{cases} \frac{r}{k_s^2 - k_n^2} [k_s B_1(k_s r) B_0(k_n r) - k_n B_0(k_s r) B_1(k_n r)], & k_s \neq k_n \\ \frac{r^2}{2} [B_0^2(k_s r) + B_1^2(k_s r)], & k_s = k_n \end{cases}, \quad (\text{C1})$$

the integrals in Eqs. (28)–(33) can be obtained as

$$\langle \psi_{A,n} \psi_{A,s} \rangle_{0 \leq r \leq r_1} = \begin{cases} \frac{r_1^2}{2} J_0^2(k_{r,A,n} r_1) & n = s, \\ 0 & n \neq s, \end{cases} \quad (\text{C2})$$

$$\langle \psi_{B,n} \psi_{B,s} \rangle_{r_1 \leq r \leq r_2} = \begin{cases} \frac{1}{2}(r_2^2 - r_1^2) & \text{for } n = s = 0, \\ \frac{1}{2}[r_2^2 \psi_{B,s}^2(r_2) - r_1^2 \psi_{B,s}^2(r_1)] & \text{for } n = s = 1, 2, \dots, \\ 0 & \text{for } n \neq s, \end{cases} \quad (\text{C3})$$

$$\langle \psi_{C,n,u_x} \psi_{C,s,u_x} \rangle_{0 \leq r \leq r_2} = \begin{cases} \frac{r_1^2}{2} J_0^2(k_{r,C,s} r_1) + \frac{r_2^2}{2} \left(\frac{\rho_0}{\bar{\rho}} \right)^2 \{ [C_3 J_0(\tilde{k}_{r,C,s} r_2) + C_4 Y_0(\tilde{k}_{r,C,s} r_2)]^2 + [C_3 J_1(\tilde{k}_{r,C,s} r_2) + C_4 Y_1(\tilde{k}_{r,C,s} r_2)]^2 \} \\ - \frac{r_1^2}{2} \left(\frac{\rho_0}{\bar{\rho}} \right)^2 \{ [C_3 J_0(\tilde{k}_{r,C,s} r_1) + C_4 Y_0(\tilde{k}_{r,C,s} r_1)]^2 + [C_3 J_1(\tilde{k}_{r,C,s} r_1) + C_4 Y_1(\tilde{k}_{r,C,s} r_1)]^2 \} & \text{for } n = s, \\ 0 & \text{for } n \neq s, \end{cases} \quad (\text{C4})$$

$$\langle \psi_{D,n} \psi_{D,s} \rangle_{r_1 \leq r \leq r_2} = \begin{cases} \frac{1}{2}(r_2^2 - r_1^2) & \text{for } n = s = 0, \\ \frac{1}{2}[r_2^2 \psi_{D,s}^2(r_2) - r_1^2 \psi_{D,s}^2(r_1)] & \text{for } n = s = 1, 2, \dots, \\ 0 & \text{for } n \neq s, \end{cases} \quad (\text{C5})$$

$$\langle \psi_{E,n} \psi_{E,s} \rangle_{0 \leq r \leq r_1} = \begin{cases} \frac{r_1^2}{2} J_0^2(k_{r,E,n} r_1) & \text{for } n = s, \\ 0 & \text{for } n \neq s, \end{cases} \quad (\text{C6})$$

$$\langle \psi_{C,n,u_x} \psi_{A,s} \rangle_{0 \leq r \leq r_1} = \begin{cases} r_1^2 [J_0^2(k_{r,A,s} r_1) + J_1^2(k_{r,C,n} r_1)], & \text{for } k_{r,A,s} = k_{r,C,n} \\ \frac{r_1 [k_{r,C,n} J_1(k_{r,C,n} r_1) J_0(k_{r,A,s} r_1) - k_{r,A,s} J_0(k_{r,C,n} r_1) J_1(k_{r,A,s} r_1)]}{k_{r,C,n}^2 - k_{r,A,s}^2}, & \text{for } k_{r,A,s} \neq k_{r,C,n} \end{cases}, \quad (\text{C7})$$

$$\begin{aligned}
& \langle \psi_{C,n,u_x} \psi_{B,s} \rangle_{r_1 \leq r \leq r_2} \\
& = C_3 \frac{\rho_0}{\bar{\rho}} \left\{ \begin{aligned} & \frac{r_2^2}{2} \left[J_0(\bar{k}_{r,C,n} r_2) - \frac{J_1(\bar{k}_{r,C,n} r_2)}{Y_1(\bar{k}_{r,C,n} r_2)} Y_0(\bar{k}_{r,C,n} r_2) \right]^2 - \frac{r_1^2}{2} \left[J_0(\bar{k}_{r,C,n} r_1) - \frac{J_1(\bar{k}_{r,C,n} r_1)}{Y_1(\bar{k}_{r,C,n} r_1)} Y_0(\bar{k}_{r,C,n} r_1) \right]^2, \quad \bar{k}_{r,C,n} = k_{r,B,s} \\ & \frac{r_2 \bar{k}_{r,C,n}}{\bar{k}_{r,C,n}^2 - k_{r,B,s}^2} \left[J_1(\bar{k}_{r,C,n} r_2) + \frac{C_4}{C_3} Y_1(\bar{k}_{r,C,n} r_2) \right] \left[J_0(k_{r,B,s} r_2) - \frac{J_1(k_{r,B,s} r_2)}{Y_1(k_{r,B,s} r_2)} Y_0(k_{r,B,s} r_2) \right] \\ & - \frac{r_1 \bar{k}_{r,C,n}}{\bar{k}_{r,C,n}^2 - k_{r,B,s}^2} \left[J_1(\bar{k}_{r,C,n} r_1) + \frac{C_4}{C_3} Y_1(\bar{k}_{r,C,n} r_1) \right] \left[J_0(k_{r,B,s} r_1) - \frac{J_1(k_{r,B,s} r_1)}{Y_1(k_{r,B,s} r_1)} Y_0(k_{r,B,s} r_1) \right], \quad \bar{k}_{r,C,n} \neq k_{r,B,s} \\ & + \frac{r_1 k_{r,B,s}}{\bar{k}_{r,C,n}^2 - k_{r,B,s}^2} \left[J_0(\bar{k}_{r,C,n} r_1) + \frac{C_4}{C_3} Y_0(\bar{k}_{r,C,n} r_1) \right] \left[J_1(k_{r,B,s} r_1) - \frac{J_1(k_{r,B,s} r_1)}{Y_1(k_{r,B,s} r_1)} Y_0(k_{r,B,s} r_1) \right] \end{aligned} \right. \quad (C8)
\end{aligned}$$

$$\begin{aligned}
& \langle \psi_{C,n,u_x} \psi_{D,s} \rangle_{r_1 \leq r \leq r_2} \\
& = C_3 \frac{\rho_0}{\bar{\rho}} \left\{ \begin{aligned} & \frac{r_2^2}{2} \left[J_0(\bar{k}_{r,C,n} r_2) - \frac{J_1(\bar{k}_{r,C,n} r_2)}{Y_1(\bar{k}_{r,C,n} r_2)} Y_0(\bar{k}_{r,C,n} r_2) \right]^2 - \frac{r_1^2}{2} \left[J_0(\bar{k}_{r,C,n} r_1) - \frac{J_1(\bar{k}_{r,C,n} r_1)}{Y_1(\bar{k}_{r,C,n} r_1)} Y_0(\bar{k}_{r,C,n} r_1) \right]^2, \quad \bar{k}_{r,C,n} = k_{r,D,s} \\ & \frac{r_2 \bar{k}_{r,C,n}}{\bar{k}_{r,C,n}^2 - k_{r,D,s}^2} \left[J_1(\bar{k}_{r,C,n} r_2) + \frac{C_4}{C_3} Y_1(\bar{k}_{r,C,n} r_2) \right] \left[J_0(k_{r,D,s} r_2) - \frac{J_1(k_{r,D,s} r_2)}{Y_1(k_{r,D,s} r_2)} Y_0(k_{r,D,s} r_2) \right] \\ & - \frac{r_1 \bar{k}_{r,C,n}}{\bar{k}_{r,C,n}^2 - k_{r,D,s}^2} \left[J_1(\bar{k}_{r,C,n} r_1) + \frac{C_4}{C_3} Y_1(\bar{k}_{r,C,n} r_1) \right] \left[J_0(k_{r,D,s} r_1) - \frac{J_1(k_{r,D,s} r_1)}{Y_1(k_{r,D,s} r_1)} Y_0(k_{r,D,s} r_1) \right], \quad \bar{k}_{r,C,n} \neq k_{r,D,s} \\ & + \frac{r_1 k_{r,D,s}}{\bar{k}_{r,C,n}^2 - k_{r,D,s}^2} \left[J_0(\bar{k}_{r,C,n} r_1) + \frac{C_4}{C_3} Y_0(\bar{k}_{r,C,n} r_1) \right] \left[J_1(k_{r,D,s} r_1) - \frac{J_1(k_{r,D,s} r_1)}{Y_1(k_{r,D,s} r_1)} Y_0(k_{r,D,s} r_1) \right] \end{aligned} \right. \quad (C9)
\end{aligned}$$

$$\langle \psi_{C,n,u_x} \psi_{E,s} \rangle_{0 \leq r \leq r_1} = \begin{cases} r_1^2 [J_1^2(k_{r,C,n} r_1) + J_0^2(k_{r,E,s} r_1)], & \text{for } k_{r,E,s} = k_{r,C,n}, \\ \frac{r_1 [k_{r,C,n} J_1(k_{r,C,n} r_1) J_0(k_{r,E,s} r_1) - k_{r,A,s} J_0(k_{r,C,n} r_1) J_1(k_{r,E,s} r_1)]}{k_{r,C,n}^2 - k_{r,E,s}^2}, & \text{for } k_{r,E,s} \neq k_{r,C,n}. \end{cases} \quad (C10)$$

¹K. S. Peat and K. L. Rathi, "A finite element analysis of the convected acoustic wave motion in dissipative silencers," *J. Sound Vib.* **184**, 529–545 (1995).

²A. Cummings and R. J. Astley, "Finite element computation of attenuation in bar silencers and comparison with measured data," *J. Sound Vib.* **196**, 351–369 (1996).

³SYSNOISE users' manual, Rev. 5.5, LMS International, 2001.

⁴A. F. Seybert, R. A. Seman, and M. D. Lattuca, "Boundary element prediction of sound propagation in ducts containing bulk absorbing materials," *J. Vib. Acoust.* **120**, 976–981 (1998).

⁵A. Selamet, I. J. Lee, and N. T. Huff, "Acoustic attenuation of hybrid silencers," *J. Sound Vib.* **262**, 509–527 (2003).

⁶R. J. Astley and A. Cummings, "A finite element scheme for attenuation in ducts lined with porous material: Comparison with experiment," *J. Sound Vib.* **116**, 239–263 (1987).

⁷A. Cummings, "A segmented Rayleigh–Ritz method for predicting sound transmission in a dissipative exhaust silencer of arbitrary cross section," *J. Sound Vib.* **187**, 23–37 (1995).

⁸R. Glav, "The point-matching method on dissipative silencers of arbitrary cross section," *J. Sound Vib.* **189**, 123–135 (1996).

⁹R. Glav, "The null-field approach to dissipative silencers of arbitrary cross section," *J. Sound Vib.* **189**, 489–509 (1996).

¹⁰R. Glav, "The transfer matrix for a dissipative silencer of arbitrary cross section," *J. Sound Vib.* **236**, 575–594 (2000).

¹¹R. Kirby, "Transmission loss prediction for dissipative silencers of arbitrary cross section in the presence of mean flow," *J. Acoust. Soc. Am.* **114**, 200–209 (2003).

ary cross section in the presence of mean flow," *J. Acoust. Soc. Am.* **114**, 200–209 (2003).

¹²A. Cummings and I. J. Chang, "Sound attenuation of a finite length dissipative flow duct silencer with internal mean flow in the absorbent," *J. Sound Vib.* **127**, 1–17 (1988).

¹³R. Kirby, "Simplified techniques for predicting the transmission loss of a circular dissipative silencer," *J. Sound Vib.* **243**, 403–426 (2001).

¹⁴M. B. Xu, A. Selamet, I. J. Lee, and N. T. Huff, "Sound attenuation in dissipative expansion chambers," *J. Sound Vib.* **272**, 1125–1133 (2004).

¹⁵A. Selamet, M. B. Xu, I. J. Lee, and N. T. Huff, "Analytical approach for sound attenuation in perforated dissipative silencers," *J. Acoust. Soc. Am.* **115**, 2091–2099 (2004).

¹⁶M. B. Xu, X. M. Zhang, and W. H. Zhang, "The effect of wall joint on the vibrational power flow propagation in a fluid-filled shell," *J. Sound Vib.* **224**, 395–410 (1999).

¹⁷M. Abom, "Derivation of four-pole parameters including higher order mode effects for expansion chamber mufflers with extended inlet and outlet," *J. Sound Vib.* **137**, 403–418 (1990).

¹⁸A. Selamet and Z. L. Ji, "Acoustic attenuation performance of circular expansion chambers with extended inlet/outlet," *J. Sound Vib.* **223**, 197–212 (1999).

¹⁹A. Selamet and Z. L. Ji, "Case study: Acoustic attenuation performance of circular expansion chambers with extended end-inlet and side-outlet," *Noise Control Eng. J.* **48**, 60–66 (2000).

- ²⁰ A. Selamet, F. D. Denia, and A. J. Besa, "Acoustic behavior of circular dual-chamber mufflers," *J. Sound Vib.* **265**, 967–985 (2003).
- ²¹ A. J. Torregrosa, A. Broatch, R. Payri, and F. Gonzalez, "Numerical estimation of end corrections in extended-duct and perforated-duct mufflers," *J. Vibr. Acoust.* **121**, 302–308 (1999).
- ²² J. W. Sullivan and M. J. Crocker, "Analysis of concentric-tube resonators having unpartitioned cavities," *J. Acoust. Soc. Am.* **64**, 207–215 (1978).
- ²³ J. W. Sullivan, "A method for modeling perforated tube muffler components. I. Theory," *J. Acoust. Soc. Am.* **66**, 772–778 (1979).
- ²⁴ J. W. Sullivan, "A method for modeling perforated tube muffler components. II. Applications," *J. Acoust. Soc. Am.* **66**, 779–788 (1979).

Absorptive properties of rigid porous media: Application to face centered cubic sphere packing

S. Gasser^{a)} and F. Paun

Office National d'Etudes et de Recherches Aérospatiales, Département Matériaux Métalliques et Procédés,
29 avenue de la division Leclerc, 92322 Châtillon cedex, France

Y. Bréchet

Institut National Polytechnique de Grenoble, Laboratoire de Thermodynamique et Physico-Chimie
Métallurgiques, Domaine universitaire, BP 75, 38402 Saint-Martin d'Hères cedex, France

(Received 29 February 2004; revised 8 January 2005; accepted 10 January 2005)

The classical description of porous media as a homogeneous equivalent fluid is presented, and its foundations on the homogenization method is introduced and applied to the numerical prediction model for a periodic porous medium composed by face centered cubic sphere packing on which measurements have been made. The results are compared with existing numerical results in the literature and with new and experimental data. © 2005 Acoustical Society of America.

[DOI: 10.1121/1.1863052]

PACS numbers: 43.55.Ev, 43.20.Bi [ANN]

Pages: 2090–2099

I. INTRODUCTION

Noise reduction is currently a major issue, especially in aeronautics. One possible way to handle the problem among several others such as reducing noise production by engines or enhancing the design of gears, and other mechanical components, is to absorb sound close to where it is produced (e.g., inside the turboengines of aircraft). This approach requires the development of acoustic liners made of materials withstanding temperatures (up to 1000 K).¹

As polymers are unsuitable for those temperatures, the remaining material classes are metals and ceramics. However, these materials are rigid with very low intrinsic damping factors. Thus, the only way to absorb sound with them is to design a porous structure.

In this paper we deal with a particular choice of porous structure, face centered cubic (FCC) packing of spheres, as shown in Fig. 1. However, most of the methods presented in this paper can be applied to other rigid porous materials with an open porosity like metal foams and felts and particularly to porous media with a periodic structure for which the acoustical properties can be computed.

In this paper we begin with an introduction to the classical modeling of porous medium acoustics (Sec. II) followed by homogenization theory, which is usually presented as a rigorous derivation of the hypotheses required by the classical models (Sec. III), and is applied (Sec. IV) to predict the acoustic properties of a FCC packing of spheres via a numerical method. Comparisons are made with results available in the literature.

II. CLASSICAL APPROACH

Since the seminal paper of Kirchhoff,² the underlying physics of the phenomena of sound propagation and absorption in a porous material has been well known. It remains the

basis for the phenomenological models to which numerical calculations can be compared. The hypothesis and the methods used in the classical approach are recalled in Sec. II A. An applied review of the models available in the literature for different hypothesis on the structure of the porous media is given in Sec. II B which allows the presentation of a general model including all those existing as special cases.

A. Fundamental hypotheses and method in the classical approach

In Kirchhoff's model for the propagation of sound in a tube, the air, carrying the acoustic wave is a viscous, thermally conducting, Newtonian fluid, and its thermodynamic properties are well described as an ideal gas. The solid fraction of the porous material is assumed rigid and at a uniform, constant temperature. These hypotheses are valid for common cellular materials made of metal or ceramics for which the stiffness of the solid skeleton is much higher than the bulk modulus of the air and the heat capacity of the solid fraction is much higher than that of the air.

In the present work, consideration is restricted to the absorptive behavior of the porous material at equilibrium, i.e., at room temperature and without any main gas stream. Moreover, the acoustic waves are assumed to be of low amplitude to remain in the linear acoustics regime. It is thus easy to work in the harmonic regime by applying a Fourier time transform.

The previous assumptions enable focus on a precise set of equations and boundary conditions. The density, temperature and pressure of the air at rest are denoted, respectively, by ρ_0 , T_0 and P_0 . The use of complex notations for harmonic waves of frequency, f , is retained with the $e^{i\omega t}$ time dependence, where $\omega = 2\pi f$ is the angular frequency. The complex amplitudes of the air speed and the fluctuating parts of the density, temperature and pressure of the air are, respectively, denoted by \mathbf{u} , ξ , τ and p .

The specific heats at constant volume, C_V , and at constant pressure, C_P , and their ratio $\gamma = C_P/C_V$ will be used as

^{a)}Current affiliation: SEGIME, 8 rue de l'hôtel de ville, 92200 Neuilly sur Seine, France.

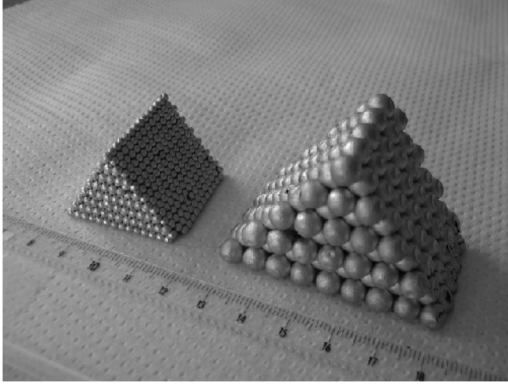


FIG. 1. Two prismatic samples of FCC nickel hollow spheres packings (2 mm diameter beads on the left, 6 mm diameter beads on the right).

well as the dynamic and bulk viscosities, μ and λ , the heat conductivity, K , and the Prandtl number, $\text{Pr} = \mu C_P / K$.

Air is assumed to follow the equation of state of a perfect gas:

$$\frac{P_0}{\rho_0 T_0} = C_P - C_V. \quad (1)$$

By the use of logarithmic derivatives, one also gets a relationship between fluctuating quantities:

$$\frac{p}{P_0} - \frac{\tau}{T_0} - \frac{\xi}{\rho_0} = 0. \quad (2)$$

The dynamics of the fluid is governed by the Navier–Stokes equation,³ which can be rewritten after linearization and time Fourier transform as

$$\rho_0 i \omega \mathbf{u} = -\mathbf{grad} p + (\lambda + \mu) \mathbf{grad} \text{div} \mathbf{u} + \mu \Delta \mathbf{u}. \quad (3)$$

Equation (3) must be completed by boundary conditions. As the solid fraction of the porous material is rigid, it has no relative motion and the no-slip condition at the solid interface Σ is written as

$$\mathbf{u}|_{\Sigma} = \mathbf{0} \quad (4)$$

The two previous equations are not sufficient to completely define the fluid flow; the continuity equation³ must be included in the discussion. After linearization and Fourier time transform, it becomes, after Eq. (2), used to eliminate ξ :

$$\text{div} \mathbf{u} + i \omega \frac{\xi}{\rho_0} = \text{div} \mathbf{u} + i \omega \left(\frac{p}{P_0} - \frac{\tau}{T_0} \right) = 0. \quad (5)$$

As Eq. (5) includes the fluctuating temperature, there is a coupling between the dynamics of the acoustic wave and its thermal behavior. It was therefore necessary to include a heat equation in the problem.

The heat equation, linearized and Fourier time transformed, is

$$\rho_0 i \omega C_V \tau = -P_0 \text{div} \mathbf{u} + K \Delta \tau. \quad (6)$$

Equation (6) can be slightly transformed by the use of (5) and (1) to obtain

$$\rho_0 i \omega C_P \tau = i \omega p + K \Delta \tau. \quad (7)$$

As for (3), Eq. (7) needs the definition of boundary conditions. From the assumption of the isothermal solid fraction, one gets

$$\tau|_{\Sigma} = 0, \quad (8)$$

by assuming the continuity of the temperature at the solid interface, which is commonly valid if the fluid is not rarefied (see Bruneau *et al.*⁴ for more general boundary conditions).

The set of equations, (3), (4), (5), (7) and (8), together with the geometry of the pores in the material, define the general mathematical problem to be solved. In principle, any problem of linear acoustics involving porous materials could be dealt within this formalism. However, it cannot be handled efficiently under this formulation except for the case of very simple pore geometries such as Kirchhoff's cylinder.² Therefore, in order to proceed further it is necessary to introduce a simpler description of the acoustics of porous media.

An additional hypothesis of local incompressibility greatly simplifies the problem. It is valid for small Mach numbers and when the local length scale is negligible compared to the acoustic wavelength, two hypotheses which are fulfilled in the linear acoustic regime considered and range of frequency of interest. This local incompressibility hypothesis leads to the concepts of effective density and effective compressibility.

Incompressibility at local scale means that the flow velocity has a zero divergence. Equation (5) is now transformed into

$$\text{div} \mathbf{u} = 0, \quad (9)$$

and Eq. (3) becomes

$$\rho_0 i \omega \mathbf{u} = -\mathbf{grad} p + \mu \Delta \mathbf{u}. \quad (10)$$

Equations (9) and (10) with their boundary condition (4) are now fully decoupled from Eq. (7) and its boundary condition (8). The first set of equations defines a dynamic problem, whereas the second defines a thermal one.

Consider first the dynamic problem: at a given frequency f , the flow is generated by a fluctuating pressure gradient. This pressure gradient $\mathbf{grad} p$ can be separated in a locally constant vector $\mathbf{g} = \mathbf{grad} \langle \mathbf{g} \cdot \mathbf{x} \rangle$ and a corrective pressure gradient $\mathbf{grad} q$. The vector \mathbf{g} , which therefore acts as a macroscopic pressure gradient, will be denoted in the following by the more explicit expression $\mathbf{grad} \langle p \rangle$, where $\langle \cdot \rangle$ represents the local spatial average of a quantity in the pore's domain. By linearity of the dynamic problem, there exists a linear relationship between the constant vector \mathbf{g} and the corresponding velocity field \mathbf{u} , and thus between \mathbf{g} and the average velocity $\langle \mathbf{u} \rangle$. This linear relationship between two vectors imposes the introduction of a second order tensor, the effective density tensor, ρ_{eff} :

$$\rho_{\text{eff}} i \omega \langle \mathbf{u} \rangle = -\mathbf{grad} \langle p \rangle. \quad (11)$$

Notice that (11) corresponds formally to an Euler equation. The dissipative behavior is now included in the complex character of the effective density instead of the viscosity.

If the porous medium is isotropic or presents cubic symmetry, however, the effective density tensor reduces to a sca-

lar effective density ρ_{eff} . For the porous materials commonly used in acoustics, this assumption is satisfactory, and the FCC packing of spheres considered here is, by definition, of cubic symmetry: the consequences of anisotropy will not be discussed in the sequel.⁵

The thermal problem can be treated in the same spirit: here, the temperature field is driven by the fluctuating pressure p , which is nearly homogeneous at a local scale. It is then convenient to consider the pressure p as equal to its local average $\langle p \rangle$, and by linearity of the thermal problem, the temperature field τ and its average $\langle \tau \rangle$ are proportional to $\langle p \rangle$. This yields the introduction of a scalar effective compressibility χ_{eff} :

$$\frac{\langle \tau \rangle}{T_0} = \left(\frac{1}{P_0} - \chi_{\text{eff}} \right) \langle p \rangle. \quad (12)$$

The meaning of this complex quantity appears when (11) and (12) are recoupled, this time at a macroscopic level. From the original continuity equation (5), one gets a macroscopic relation by averaging all quantities at a local scale:

$$\langle \text{div } \mathbf{u} \rangle + i\omega \left(\frac{\langle p \rangle}{P_0} - \frac{\langle \tau \rangle}{T_0} \right) = 0. \quad (13)$$

The first term of (13) is the mean relative increase of volume of a fluid particle, and as the flow is assumed incompressible at a local scale, it reduces to its macroscopic contribution $\text{div} \langle \mathbf{u} \rangle$. With (12), Eq. (13) thus becomes

$$\text{div} \langle \mathbf{u} \rangle = -i\omega \chi_{\text{eff}} \langle p \rangle, \quad (14)$$

and χ_{eff} corresponds formally to the compressibility of the porous medium. Finally, (11) and (14) allows the derivation of a dissipative wave equation in the harmonic regime:

$$\Delta \langle p \rangle + \rho_{\text{eff}} \chi_{\text{eff}} \omega^2 \langle p \rangle = 0. \quad (15)$$

The previous equations and boundary conditions constitute the mathematical formulation of the problem, which has to be solved for the geometry of the selected porous medium.

B. Application to model porous media proposed in the literature

Almost all of the theoretical models proposed in the literature to describe sound propagation and absorption in rigid porous media are successive improvements on the analytical solution for a simple porous medium made of cylindrical tubes. Thus, they can be also considered as simplified versions of a more general solution, which is why the most general model has been presented. However, it is worth recalling the steps that led to this general model.

Kirchhoff's paper¹ dealt with the propagation of an acoustic wave in a cylinder. From his results, a very simple model of porous medium can be built consisting of a collection of identical parallel straight tubes of circular cross-section. The only parameters that this model requires are the radius of the cylinders and the fraction of cylinders in a volume of porous material, i.e. its porosity φ .

As pores are hardly straight in any realistic material, a further step is achieved by considering tortuous tubes. Several authors⁶⁻⁸ introduced a scalar parameter to take the effect of this sinuosity into account with different conventions

for each author; the present paper considers the parameter called tortuosity after Johnson *et al.*,⁸ denoted by α_∞ as it is related to the infinite frequency limit of the dynamic problem. For thin tortuous tubes of constant cross-section, it is equal to $(L/L_0)^2$, where L is the length of the tortuous tube and L_0 is the straight line linking both ends of the tube.

Another generalization concerns the shape of the cross section: if the tubes are no longer circular, two different equivalent radii can be introduced, each corresponding to one end of the frequency spectrum. At zero frequency, the flow is completely viscous and thus follows Darcy's law:

$$\varphi \langle \mathbf{u} \rangle = -\frac{k_0}{\mu} \text{grad} \langle p \rangle, \quad (16)$$

where k_0 is permeability. For circular cylinders of radius R , it corresponds to a Poiseuille flow and $k_0 = \varphi R^2/8$, where φ is the porosity of the porous material. More generally, for any geometry, a low frequency length l can be introduced, defined by $k_0 = \varphi l^2/8$.

In contrast, for very high frequencies, the viscosity has an effect in the vicinity of the solid, in a layer that gets thinner with an increasing frequency. The infinite frequency limit is an inviscid incompressible flow, and a characteristic length Λ , playing the role of a high frequency radius, as introduced by Johnson *et al.*,⁸ defined by

$$\frac{2}{\Lambda} = \frac{\int \int_{\Sigma} \mathbf{u}^2 dS}{\int \int_{\Omega} \mathbf{u}^2 dV}, \quad (17)$$

where Ω is the fluid domain and Σ is its solid border. It is customary to work with Λ and with a shape factor M defined by $l^2 = M \Lambda^2 / \alpha_\infty$ instead of l .

When the cross-section does not vary along the tube, and when its orientation varies smoothly, the speed is always nearly axial and the dynamic problem loses its vector character. Thus, the thermal problem and the dynamic problem have the same analytic form and the geometrical lengths R and Λ are common to both problems. The next generalization comes when the porous medium to be considered does not look like such tubes. Each problem thus has its own characteristic lengths, and two quantities Λ' and M' , respectively, a thermal characteristic length and a thermal form factor, have to be introduced.^{9,10} At this point, the most general model may be found in Allard's classical monograph;¹¹ additionally, a last set of parameters, a dynamic parameter, p , and a thermal parameter, p' , have been introduced to improve the models at low frequency^{12,13} to let the theoretical expressions for the effective density and compressibility match the numerical values at the second order in ω . To the authors' knowledge, they do not have any physical interpretation. Such quantities cannot be mixed up with the pressure, so these notations are kept in this paper.

All of these parameters (porosity φ , tortuosity α_∞ , characteristic lengths Λ and Λ' , shape factors M and M' , supplementary parameters p and p') correspond to either the low or high frequency limit. The general expression to interpolate to finite frequency relies on a remark by Biot⁶ that for straight tubes of several different cross-sections, the corresponding effective densities have the same shape. The theoretical ex-

planation was again provided by Johnson *et al.*⁸ It is based on the principle of causality which restricts the possible interpolating functions for both the effective compressibility and density to have their zeros and poles on the negative portion of the imaginary axis (Champoux and Allard⁹).

$$\rho_{\text{eff}} = \rho_0 \alpha_\infty \left(1 - i \frac{8\mu}{\rho_0 \Lambda^2} \frac{1 - p + p \left(1 + \frac{M^2 \rho_0 \Lambda^2}{2p^2} i \omega \right)^{1/2}}{M \omega} \right), \quad (18)$$

and Lafarge's effective compressibility is

$$\chi_{\text{eff}} = \frac{1}{\gamma P_0} \left(\gamma - \frac{\gamma - 1}{1 - i \frac{8\mu}{\rho_0 \text{Pr} \Lambda'^2} \frac{1 - p' + p' \left(1 + \frac{M'^2 \rho_0 \Lambda'^2 \text{Pr}}{2p'^2} i \omega \right)^{1/2}}{M' \omega}} \right). \quad (19)$$

It must be stressed that all the models presented above make use at least implicitly of the local incompressibility hypothesis. In the next section, the homogenization technique is introduced for two reasons: first, to give a mathematically coherent derivation of the incompressibility hypothesis and of the introduction of the effective density and compressibility, and second to obtain an effective numerical scheme for computing those effective quantities. As far as the general model derivation (and all the simpler models in the literature) does not make use of homogenization, but relies on assumptions which can be derived by homogenization, it makes sense to compare numerical computations based on homogenization with theoretical results using the general model.

III. HOMOGENIZATION

The previous derivation (Sec. II) is not completely rigorous. For example, the incompressibility condition (9) is used with the dynamics equation (3), but it should not be used with the heat equation (6). Another important point is that the set of PDEs (3), (4), (5), (7) and (8) do not define a well-posed problem as the boundary conditions (4) and (8) are defined on the solid interface Σ only. As it was assumed that the fluid percolates in the porous medium, it is not surrounded by the solid, so the limits of the fluid domain are not precisely defined and there is a lack of boundary conditions to complete the PDEs. Finally, all the geometries on which the models were derived have a tubular structure, and the tubes do not form a network as the pores of an open porous medium would. These geometries are also basically anisotropic as the all tubes are directed along a same direction, while the corresponding models intend to describe isotropic media.

When the porous medium has a periodic structure, a mathematical technique, known as periodic homogenization,¹⁴ gives a strong basis for the theory of the

The general model, proposed by Pride *et al.*¹² for the effective density on one hand, and by Lafarge¹³ for the effective compressibility on the other, corresponds to an arbitrary choice of functions parametrized by φ , α_∞ , Λ , Λ' , M , M' , p and p' . The Pride *et al.* effective density is given by

previous paragraph; it ensures that the derivation of Eqs. (11) and (14) is coherent, and it provides a way to compute the effective quantities ρ_{eff} and χ_{eff} . As this paper focuses on predicting the acoustical behavior of a FCC packing of spheres, the technique developed later is suitable.¹⁵

First, the standard homogenization scheme is introduced (Sec. III A), leading to its numerical application using the finite element method (Sec. III B).

A. Standard homogenization scheme

The periodic homogenization technique and its results are now briefly described. The presentation of Boutin *et al.*¹⁴ is followed here and we refer to that paper for the details of their derivation in order to focus on the principal hypotheses and results.

The homogenization procedure begins by introducing a ratio ϵ of a characteristic length of the local scale to a characteristic length of the macroscopic scale. When considering a periodic structure, the natural choice of the small length scale is the size of the unit cell, whereas the natural macroscopic length of an acoustical phenomenon is its wavelength. Homogenized equations are expected to hold when this ratio ϵ is small enough.

The procedure can be split into two steps: the first is physical, and the second is mathematical. The physical step is a matter of estimates, and introducing dimensionless quantities may help. Let us therefore make all the variables dimensionless. For example, the reference pressure being the ambient pressure, P_0 , the acoustic pressure, p , is replaced by a reduced pressure, $p' = p/P_0$. The other reference quantities that can be chosen are the angular frequency, ω , the ambient temperature, T_0 , the isothermal celerity, $c_0 = (P_0/\rho_0)^{1/2}$ (the isothermal celerity is preferred to the adiabatic one to avoid unnecessary prefactors), a corresponding macroscopic length $\lambda_0 = c_0/\omega$, and a corresponding viscosity $\mu_0 = \rho_0 \omega \lambda_0^2$. For this viscosity, however, it is more convenient to chose $\mu_0 \epsilon^2$ instead. Boutin *et al.*¹⁴ argue that the characteristic length

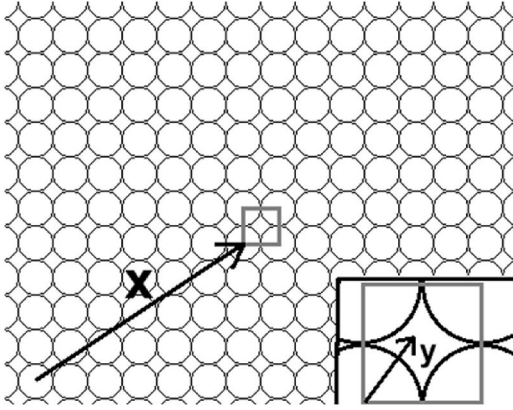


FIG. 2. The macroscopic coordinate \mathbf{x} represents the position of a unit cell, whereas the local coordinate \mathbf{y} represents the position of a point inside that cell.

corresponding to viscous effects is the local length rather than the macroscopic one, and thus an additional ϵ^2 factor coming from the second order derivatives must be introduced. That homogenization scheme, which is followed in this article, is referred to as the standard scheme.

Note also that λ_0 must be used to make the differential operators dimensionless as well as the position variable \mathbf{x} . Denoting the dimensionless variables and operators with a prime symbol, Eq. (3) becomes

$$i\mathbf{u}' = -\mathbf{grad}' p' + \epsilon^2(\lambda' + \mu')\mathbf{grad}'(\text{div}' \mathbf{u}') + \epsilon^2\mu'\Delta' \mathbf{u}'. \quad (20)$$

Equations (5) and (7) become, respectively,

$$\text{div}' \mathbf{u}' = i\tau' - ip' \quad (21)$$

and

$$i\tau' = \frac{\gamma-1}{\gamma} ip' + \frac{\mu'\epsilon^2}{\text{Pr}} \Delta' \tau', \quad (22)$$

where (1) was used and the thermal conductivity was incorporated using the Prandtl number.

The mathematical step of homogenization is actually an asymptotic method, i.e., it provides asymptotic equations from a family of equations parametrized by the ratio ϵ , when this quantity gets small. It first consists of introducing two position variables \mathbf{x} and \mathbf{y} , the former defining the position of a given point at a macroscopic scale and the latter defining it locally. To be clear, \mathbf{x} can be understood as the coordinates of the periodic cell where the point lies, and \mathbf{y} refers to its position inside the cell, as shown in Fig. 2.

As ϵ tends to zero, the cell size gets smaller and smaller, so to work on the same unit cell for any ϵ , the local coordinate, \mathbf{y} , is made dimensionless with a local length such as $\epsilon\lambda_0$, whereas the reference length for \mathbf{x} is the macroscopic one, λ_0 . The reduced coordinates, $\mathbf{x}' = \mathbf{x}/\lambda_0$ and $\mathbf{y}' = \mathbf{y}/\epsilon\lambda_0$, are thus introduced. Any physical quantity depending on position will be considered as a function of both \mathbf{x}' and \mathbf{y}' for a given ϵ . Moreover, its dependence upon ϵ is expanded as a power series as the focus is on the small ϵ limit. As an example, the speed field can be written as

$$\mathbf{u}' = \mathbf{u}'_0(\mathbf{x}', \mathbf{y}') + \epsilon \mathbf{u}'_1(\mathbf{x}', \mathbf{y}') + \epsilon^2 \mathbf{u}'_2(\mathbf{x}', \mathbf{y}') + \dots \quad (23)$$

When ϵ is small, the local behavior does not vary much from one cell to its neighbors, so the dependence upon \mathbf{y}' in (23) is also assumed to be periodic.

The homogenized field is then the average value, with respect to \mathbf{y}' , of the first order field \mathbf{u}'_0 over the periodic cell with respect to \mathbf{y}' . The interesting behavior occurs only at the macroscopic scale and higher order corrections can be ignored if ϵ is small enough. This assumption is always made when applying homogenization.^{14,16}

The dimensionless differential operators are then split into their macroscopic and local components. For example, the gradient operator is expressed as

$$\mathbf{grad}' = \mathbf{grad}'_{\mathbf{x}'} + \frac{1}{\epsilon} \mathbf{grad}'_{\mathbf{y}'}, \quad (24)$$

where the ϵ^{-1} factor comes from the different reference lengths used to make \mathbf{x} and \mathbf{y} dimensionless. When power series such as (23) and two-scale differential operators like (24) are introduced into the dimensionless equations (20), (21) and (22), it yields expansions in increasing powers of ϵ . Identifying the corresponding orders on both members of each equation thereby leads to local equations that will be made explicit.

The higher order in Eq. (21) is ϵ^{-1} , and the corresponding local equation reads as

$$\text{div}'_{\mathbf{y}'} \mathbf{u}'_0 = 0 \quad (25)$$

and the local incompressibility of the fluid speed (9) is found at first order. Equation (20) leads at first order to

$$0 = -\mathbf{grad}'_{\mathbf{y}'} p'_0, \quad (26)$$

which means that the pressure is homogeneous at first order as was assumed above for the heat equation, while at second order one obtains

$$i\mathbf{u}'_0 = -\mathbf{grad}'_{\mathbf{x}'} p'_0 - \mathbf{grad}'_{\mathbf{y}'} p'_1 + \mu' \Delta'_{\mathbf{y}'} \mathbf{u}'_0, \quad (27)$$

which is the dimensionless equation corresponding to (10) and the macroscopic gradient \mathbf{g} introduced above (Sec. II A) corresponds to the gradient of the pressure at first order with respect to the macroscopic variable. A similar conclusion can be obtained for the heat equation, showing that Eq. (7) is valid. The mathematical coherence of the classical derivation is thus justified by the homogenization technique.

The periodicity of any field with respect to \mathbf{y}' is also useful, as it completes the boundary conditions (4) and (8). Thereby, the dynamic and the thermal problems are well posed and can be solved numerically. Thanks to the superposition principle for linear equations, only one computation for the thermal problem and three for the dynamic problem (one by an independent direction of the macroscopic gradient) are required at a given frequency.

A rigorous proof of the validity of the mathematical step, especially concerning the \mathbf{y}' -periodicity and the power series expansion (23), making use of two-scale convergence, can be found in Allaire's thesis.¹⁷

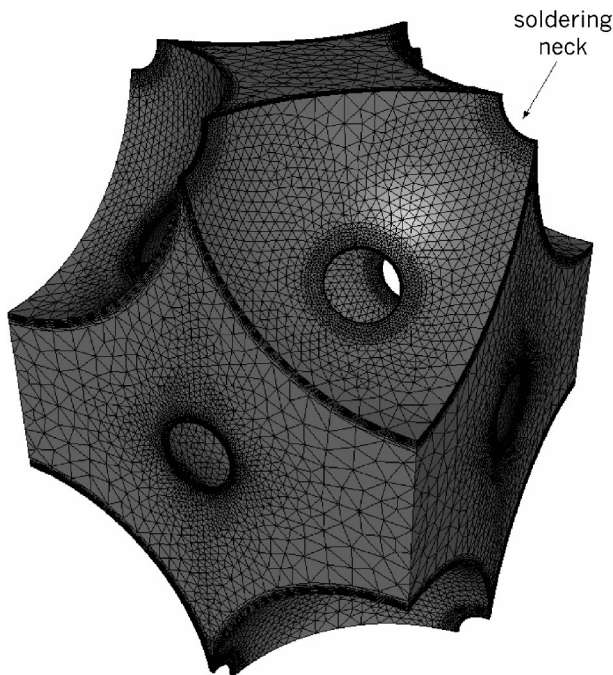


FIG. 3. Mesh of the elementary fluid cell. Anisotropic mesh refinement is visible along the border of the spheres and close to the soldering necks.

B. Numerical implementation

Now that boundary conditions for the dynamic and the thermal problems have been completed by periodic conditions owing to homogenization, both problems deserve a numerical resolution scheme; the authors chose to implement a finite-element method (FEM). Other methods have been suggested in the literature, such as the collocation method¹⁸ and the boundary element method,¹⁹ but only for the dynamic problem. In addition, the collocation method in the paper of Chapman and Higdon¹⁸ is based on spherical harmonics and it seems to be suitable for sphere-based geometries but not for more general ones.

The FEM involves a mesh of elementary polyhedra, on which any function is approximated by simple functions. For example, the present work considered a nonstructured tetrahedral mesh, with linear interpolation functions on each tetrahedron, plus a bubble function (i.e., the product of the four barycentric coordinates of the tetrahedron) for speed degrees of freedom. This last function, supplementing the set in the so-called P1-bubble/P1 approximation,²⁰ must be introduced to ensure the correct convergence of the FEM when the mesh gets thinner. The complete fluid domain of the periodic cell must be meshed. A complementary requirement is that the surface meshes on two bounding surfaces corresponding by periodicity are the same so the periodic boundary conditions can be handled.

Here a physical phenomenon already mentioned above and of capital importance for applying FEM to the problems analyzed here must be emphasized: as frequency increases, the viscous dissipation occurs closer and closer to the solid interface, in a dynamic boundary layer of thickness proportional to $\omega^{-1/2}$. Heat exchanges occur as well in a thermal boundary layer of similar thickness. To give an order of magnitude, at 1 kHz their typical thickness is about 100 μm .

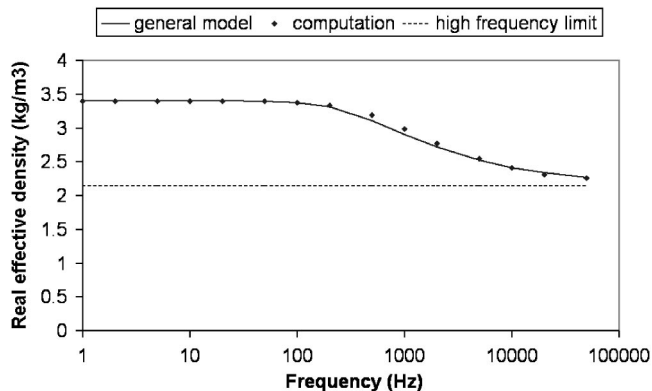


FIG. 4. Real part of the effective density.

Those boundary layers are of crucial importance to characterize absorptive properties of the porous medium at high frequency, as they concentrate the main contribution to energy dissipation. For the FEM to properly describe what happens in the vicinity of the solid, a refined mesh is thus unavoidable. However, it may be too costly in terms of memory to refine the mesh everywhere, so a local refinement close to the solid might be necessary.²¹

The resolution of the thermal problem by FEM, as it is a scalar problem with one PDE only, consists simply in solving a sparse linear system. The dynamic problem is more difficult, and Uzawa's method²² was used to apply the incompressibility condition (9) to (10). The authors refer the reader to the thesis of one of them¹⁶ for the calculation of the matrix elements for tetrahedral elements, the implementation of the boundary conditions (including periodic conditions) and the choice of the numerical method to solve a large sparse linear system being complex and symmetric (but not Hermitian).

IV. RESULTS

This strategy has been applied to a FCC stacking of beads. Figure 3 shows the mesh that was used. The radius of the spheres is here 1 mm, and a soldering neck of radius 150 μm binding the beads together was taken into account. The results of the homogenization technique can be compared to the prediction of the general model (Sec. IV A), other numerical results in the literature (Sec. IV B) and experimental results (Sec. IV C).

A. Comparison with the general model

The effective density and compressibility were computed for several frequencies from 0 to 50 kHz: the results are plotted in Figs. 4 through 6. The theoretical curves, obtained from the general model with computed parameters, are also plotted in these figures.

Concerning the effective density model, the porosity was computed from the volume of the mesh, yielding $\varphi=26\%$ as expected (the correction due to the soldering necks volume is about 0.1%); the tortuosity and the dynamic length Λ were obtained from a potential flow, leading to $\Lambda = 1.64 \times 10^{-4}$ m and $\alpha_\infty = 1.66$. From a steady flow, one can compute the permeability $k_0 = 6.83 \times 10^{-10}$ m², yielding $M = 1.29$ for the dynamic shape factor. The parameter p was obtained from a computation at a very low frequency

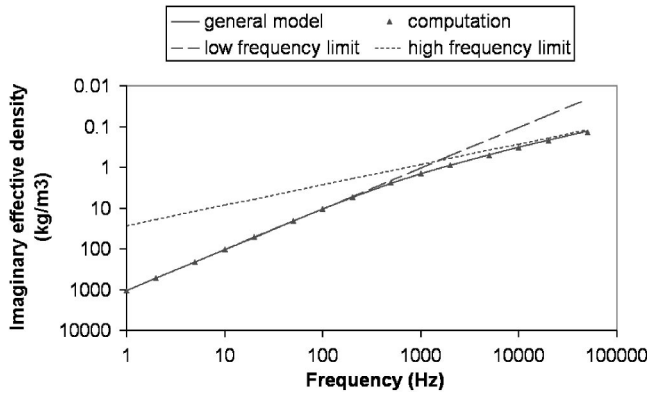


FIG. 5. Imaginary part of the effective density.

(10^{-3} Hz). The correlation between the computation and the prediction of the general model (Sec. II) is satisfactory.

For the effective compressibility, Λ' was obtained by a solid surface to volume ratio⁹ on the mesh:

$$\frac{2}{\Lambda'} = \frac{\iint_{\Sigma} dS}{\iiint_{\Omega} dV}, \quad (28)$$

yielding $\Lambda' = 2.49 \times 10^{-4}$ m. The thermal shape factor M' and the parameter p' were derived from computations at very low frequencies (10^{-3} Hz and 10^{-2} Hz) and the values $M' = 1.36$ and $p' = 0.4$ were obtained. The general model provides again a precise prediction of the computed values.

The numerical results at any frequency are correctly described by the general model. Additionally, it is worth noting that the good agreement is not the result of a parameter fitting procedure, since all the parameters were obtained from independent numerical computations at either very low frequencies or at infinite frequency. It is therefore sufficient to compute all the parameters to provide a satisfactory prediction of the effective quantities for all frequencies.

B. Comparison with numerical results of the literature

To the authors' knowledge, all the numerical work on packed beads in the literature have been performed for the dynamic problem. Chapman and Higdon¹⁸ considered the three cubic lattices (SC, BCC and FCC), with overlapping or nonoverlapping spheres depending on the prescribed porosity. In particular, they give the values of the tortuosity, of

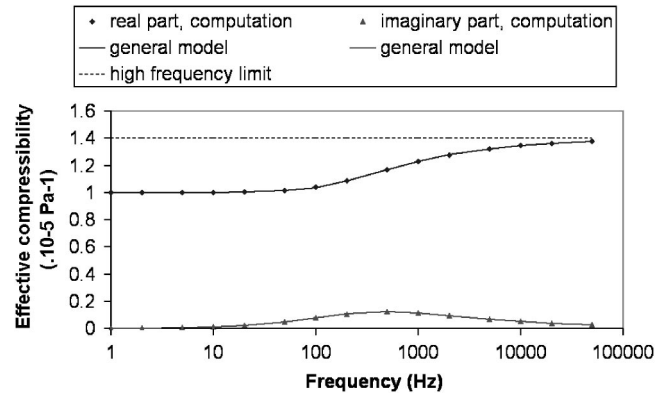


FIG. 6. Real part (above) and imaginary part (below) the effective compressibility.

Λ/R and of k_0/R^2 for spheres in contact. Zhou and Sheng²³ computed the tortuosity of a SC packing and its permeability. Finally, Maier *et al.*²⁴ computed the permeability of the FCC packing. The results are gathered in Table I. The computed value of permeability in the present work is very close to that of Maier *et al.*, whereas the Chapman and Higdon value for FCC packing is slightly superior and their tortuosity is also close to the one presented here. The characteristic length, Λ , however, is 32% lower than the value obtained in this study. A possible explanation of this disagreement is that the geometry described above takes soldering necks into account, whereas the Chapman and Higdon geometry consists in spheres with point contacts; as the boundary layer gets small for high frequencies, the influence of the local details of the geometry is stronger for physical parameters defined at high frequencies, i.e. the tortuosity and the dynamic length Λ . Tortuosity, however, usually shows small variations (it hardly exceeds 2 for common geometries), so only Λ is strongly affected by the presence of the necks.

To confirm this qualitative interpretation, a direct computation of α_∞ , Λ and Λ' was performed in the case of straight cylindrical necks (for the sake of meshing efficiency) of varying radius r . Figure 7 shows the three kinds of contact between spheres which are mentioned here (point contact, smooth neck and cylindrical neck). The values α_∞ and Λ are obtained as a result of an inviscid fluid simulation, whereas Λ' was computed as a ratio of a volume to a surface, numeri-

TABLE I. Computations of the acoustic properties of bead stackings in the literature.

Authors	Packing (porosity φ)	Tortuosity α_∞	Permeability k_0/R^2	Dynamic length Λ/R	Dynamic shape factor M
Chapman and Higdon (Ref. 17)	SC (0.48)	1.38	$1.01 \cdot 10^{-2}$	0.324	2.24
Chapman and Higdon (Ref. 17)	BCC (0.32)	1.47	$2.01 \cdot 10^{-3}$	0.192	2.00
Chapman and Higdon (Ref. 17)	FCC (0.26)	1.61	$6.95 \cdot 10^{-3}$	0.124	2.24
Zhou and Sheng (Ref. 23)	SC (0.48)	1.48	$9.84 \cdot 10^{-3}$		
Maier <i>et al.</i> (Ref. 24)	FCC (0.26)		$6.84 \cdot 10^{-4}$		
This paper	FCC (0.26)	1.66	$6.83 \cdot 10^{-4}$	0.164	1.29

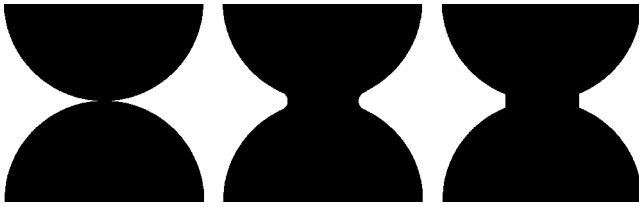


FIG. 7. Sketch of the different kinds of contact between two spheres which are considered in this work. From left to right: point contact, smooth neck and cylindrical neck.

cally from measurements on meshes and analytically, as the geometry is very simple:

$$\Lambda' = \frac{2V}{S}$$

$$= 2 \frac{(2R\sqrt{2})^3 - 4 \left[\frac{4\pi R^3}{3} - 12 \left(\pi h^2 \left(R - \frac{h}{3} \right) - \pi r^2 h \right) \right]}{4[4\pi R^2 + 12(2\pi r h - 2\pi R h)]},$$

$$h = R - \sqrt{R^2 - r^2}, \quad (29)$$

where V is the volume of the fluid domain in a conventional FCC cell, S its surface and h is the height of a half neck.²⁵

The results for Λ/R and Λ'/R are shown in Fig. 8. The computed value of Λ' is in good agreement with the analytical value and they show a substantial variation with r/R (about 20%). Moreover, the computed value of Λ' on the mesh with a smooth neck is close to the curve for cylindrical necks. Concerning Λ , the numerical results show a stronger dependence on r/R , Λ varying by a factor of 2. The value

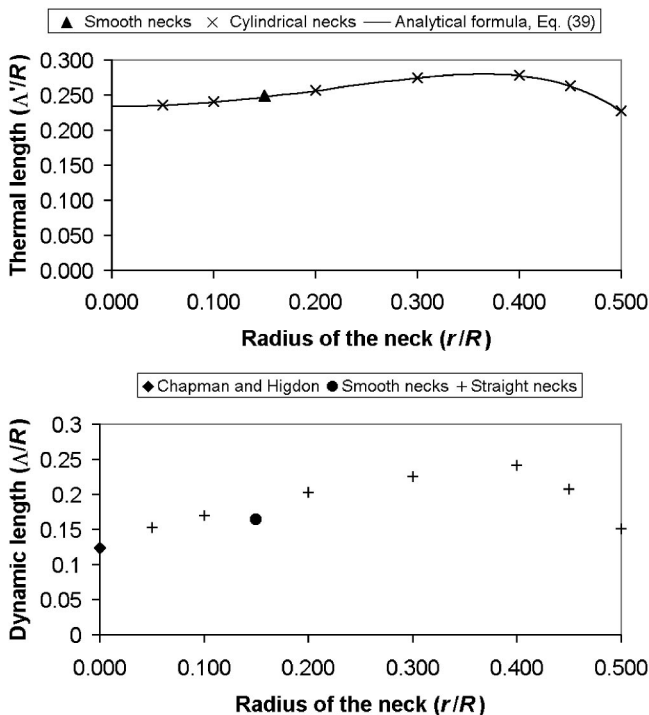


FIG. 8. Characteristic lengths of the equivalent fluid approach as a function of the normalized neck radius for different neck descriptions. (a) Thermal length Λ' ; (b) dynamic length Λ . The values computed above for smooth necks are reported on the graph, as well as the Chapman and Higdon value (Ref. 18) for Λ .

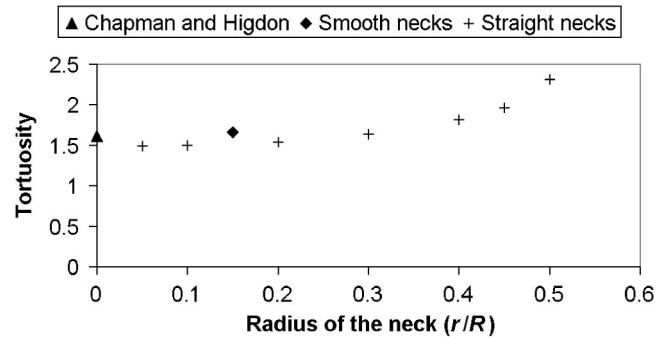


FIG. 9. Tortuosity versus neck radius for cylindrical necks. The value computed above for smooth necks and the Chapman and Higdon value are also reported on the graph.

provided by Chapman and Higdon (i.e., without a neck) seems a satisfactory extension of the computed values of Λ to the case of $r=0$ when meshing becomes difficult, and practically impossible with common meshing software. The value corresponding to the mesh with smooth necks is slightly below the Λ vs r/R curve that can be estimated from Fig. 8, indicating that the shape of the neck plays a role as well.

The numerical simulation provides also the tortuosity. Figure 9 describes its dependence on r/R . The effect of the neck size becomes relevant when r/R reaches about 0.3, the global variation of α_∞ being about 35%. The value computed by Chapman and Higdon is slightly above the plateau for small ratios r/R and the tortuosity obtained with a smooth-neck mesh is slightly above as well. The gap may come from a singular behavior of the analytical inviscid solution in the vicinity of the contact between spheres for the former, whereas for the latter, the disagreement is possibly due to the shape of the neck.

The results presented here thus may be considered to be in good agreement with previous numerical studies and it appears that the high frequency parameters show a rather strong dependence on the size of the soldering necks. It must therefore be taken into account to predict the acoustic behavior of a stack of beads on a broad range of frequencies.

C. Comparison with experimental results of the literature

The numerical results can be compared to previously published results, either experimental measurements or results from a full numerical calculation.

Several measurements of various parameters, summarized in Table II, have been performed on random packings of beads in the literature.

Allard *et al.*²⁶ is the most complete measurement, as it provides values for parameters related to the dynamic problem as well as the thermal problem. The value $M=5$, however, seems too large in comparison with other studies. Charlaix *et al.*²⁷ report measurements on three samples of 50% porosity. Kojima *et al.*^{28,29} measured several samples, two of which being made of sintered glass beads. Baker and Rudnick³⁰ also measured the acoustic parameters on sintered bronze beads. Smeulders *et al.*³¹ limited their study on high

TABLE II. Measurements of different acoustic parameters reported in the literature. The characteristic lengths Λ and Λ' are divided by the sphere radius R . Also, numbers between parentheses refer to values obtained by means of electrical resistance measurements.

Authors	Porosity φ	Tortuosity α_∞	Dynamic length Λ/R	Thermal length Λ'/R	Dynamic shape factor M	Thermal shape factor M'	Dynamic coefficient p
Allard <i>et al.</i> (Ref. 26)	0.40	1.37	0.123	0.43	5	1	
Charlaix <i>et al.</i> (Ref. 27)	0.50	1.49 (1.42)	0.527		1		0.76
Charlaix <i>et al.</i> (Ref. 27)	0.50	1.64 (1.61)	0.459		1		0.64
Charlaix <i>et al.</i> (Ref. 27)	0.50	1.56 (1.48)	0.443		0.9		0.56
Kojima <i>et al.</i> (Refs. 28, 29)	0.35	1.65	0.170		1.9		
Kojima <i>et al.</i> (Refs. 28, 29)	0.37	1.58	0.194		2.5		
Baker and Rudnick (Ref. 30)	0.43	1.61	0.33		1.34		
Baker and Rudnick (Ref. 30)	0.41	1.97	0.32		1.44		
Baker and Rudnick (Ref. 30)	0.41	2.7	0.21		3.54		
Smeulders <i>et al.</i> (Ref. 31)	0.31	2.8	0.22				
Smeulders <i>et al.</i> (Ref. 31)	0.31	2.7	0.20				
Smeulders <i>et al.</i> (Ref. 31)	0.32	2.5	0.17				
This paper	0.26	1.5	0.13	0.067			
This paper	0.36	1.5	0.13	0.2			
This paper	0.36	1.5	0.2	0.2			

frequency parameters for three samples. Finally, Berryman and Blair³² measured $k_0/R^2 = 2.84 \times 10^{-3}$ for a 30% porosity sample.

We made our own experimental measurements based on a least square fit on the absorption curve of several samples between 1000 and 5000 Hz, measured with a Kundt tube of diameter 29 mm. To reduce the number of parameters to be fitted, a simplified model was assumed where $M = M' = p = p' = 1$. The samples were 3 mm high cylinders of either

hollow nickel spheres or full steel beads, in FCC or random packing.

These experimental values provided an order of magnitude of each parameter. It does not seem possible to give a simple law relating tortuosity and porosity, but an interesting correlation can be found between the viscous characteristic length l and the porosity, as it appears in Fig. 10. From Fig. 10, it is possible to derive the following approximate linear relationship:

$$\frac{l}{R} = \frac{4}{3}(\varphi - 0.2). \quad (30)$$

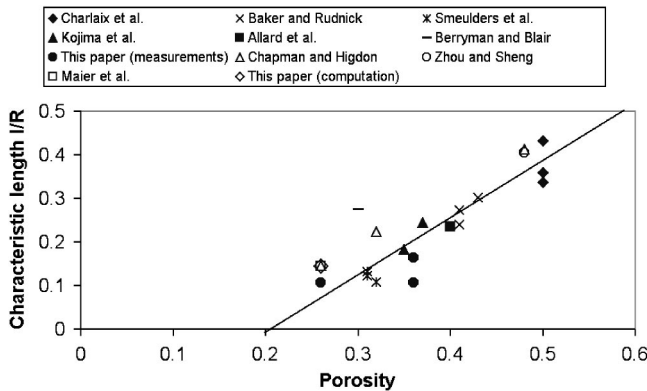


FIG. 10. The correlation between the porosity of the bead packing and the viscous characteristic length, as found in the literature. The first seven sets of values are measurements, whereas the four last sets are numerical results. The straight line shows the approximate linear relationship between the two parameters.

V. CONCLUSION

In the present contribution, the acoustic properties of a regular FCC packing of spherical hollow beads were predicted by means of homogenization and FEM computations, leading to an analytic description of that porous medium in terms of the classical equivalent fluid theory. The predictions agree with previously published numerical and experimental data.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Defense Advanced Research Project Agency (DARPA) for partial funding of this study. They also thank Denis Lafarge, Yves Aurégan, Laurent Leyeikian and Jean-Louis Auriault for very

fruitful discussions, and Jason Nadler for improving the language used in this paper. Constructive comments from referees of the paper are gratefully acknowledged.

- ¹F. Paun, S. Gasser, and L. Leylekian, "Design of materials for noise reduction in aircraft engines," *Aerospace Sci. Technol.* **7**, 63–72 (2003).
- ²G. Kirchhoff, "Über den Einfluß der Wärmeleitung in einem Gase auf die Schallbewegung," *Ann. Phys. Chem.* **134**, 177–193 (1868); translation in English: "On the influence of heat conduction in a gas on sound propagation," in *Physical Acoustics*, edited by R. B. Lindsay (Hutchison & Ross, Dowden, 1974).
- ³L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*, 2nd ed. (Butterworth-Heinemann, Woburn, MA 1987).
- ⁴M. Bruneau, P. Herzog, J. Kergomard, and J. D. Polack, "General formulation of the dispersion equation in bounded visco-thermal fluid, and application to some simple geometries," *Wave Motion* **11**, 441–451 (1989).
- ⁵Actually, there is few work on anisotropy of absorptive porous media. The interested reader may refer to M. Melon, D. Lafarge, B. Castagnède, and N. Brown, "Measurement of tortuosity of anisotropic acoustic materials," *J. Appl. Phys.* **78**, 4929–4932 (1995) as one of the few published papers on this subject.
- ⁶M. A. Biot, "Theory of propagation of elastic waves in a fluid saturated porous solid II: Higher frequency range," *J. Acoust. Soc. Am.* **28**, 179–191 (1956).
- ⁷C. Zwikker and C. W. Kosten, *Sound Absorbing Materials* (Elsevier, Amsterdam, 1949).
- ⁸D. L. Johnson, J. Koplik, and R. Dashen, "Theory of dynamic permeability and tortuosity in fluid-saturated porous media," *J. Fluid Mech.* **176**, 379–402 (1987).
- ⁹Y. Champoux and J.-F. Allard, "Dynamic tortuosity and bulk modulus in air-saturated porous media," *J. Appl. Phys.* **70**, 1975–1979 (1991).
- ¹⁰Y. Champoux and M. R. Stinson, "On acoustical models for sound propagation in rigid frame porous materials and the influence of shape factors," *J. Acoust. Soc. Am.* **92**, 1120–1131 (1992).
- ¹¹J.-F. Allard, *Propagation of Sound in Porous Media* (Elsevier Applied Science, Amsterdam, 1993).
- ¹²S. R. Pride, F. D. Morgan, and A. F. Gangi, "Drag forces of porous-medium acoustics," *Phys. Rev. B* **47**, 4964–4978 (1993).
- ¹³D. Lafarge, "Propagation du son dans les matériaux poreux à structure rigide saturés par un fluide viscothermique," Ph.D. thesis, Université du Maine, 1993; translation in English, "Sound propagation in rigid porous media saturated by a viscothermal fluid."
- ¹⁴C. Boutin, P. Royer, and J.-L. Auriault, "Acoustic absorption of porous surfacing with dual porosity," *Int. J. Solids Struct.* **35**, 4709–4737 (1998).
- ¹⁵For porous media with random structures, a stochastic homogenization technique has also been proposed in S. Torquato, *Random Heterogeneous Materials* (Springer-Verlag, Berlin, 2002) based on the hypotheses of stochastic homogeneity and ergodicity (i.e., the random porous medium is representative of a class of porous media on which a measure of probability is defined). It leads to results similar to periodic homogenization, but it is not convenient for computational purposes as it makes use of abstract probabilistic methods.
- ¹⁶St. Gasser, "Etude des propriétés acoustiques et mécaniques d'un matériau métallique poreux modèle à base de sphères creuses de nickel," Ph.D. thesis, ONERA, 2003; translation in English: "Study of the acoustical and mechanical properties of a model porous metallic material made of hollow nickel spheres."
- ¹⁷G. Allaire, "Homogénéisation des équations de Stokes et de Navier-Stokes," Ph.D. thesis, CEA, 1990; translation in English: "Homogenization of Stokes and Navier–Stokes equations."
- ¹⁸A. M. Chapman and J. J. L. Higdon, "Oscillatory Stokes flow in periodic porous media," *Phys. Fluids A* **4**, 2099–2116 (1992).
- ¹⁹L. Borne, "Harmonic Stokes flow through periodic porous media: a 3D boundary element method," *J. Comput. Phys.* **99**, 214–232 (1992).
- ²⁰F. Hecht and C. Pares, "NSPIB3: un logiciel pour résoudre les équations de Navier–Stokes incompressible 3D," INRIA Report RR-1449, 1991; translation in English: "NSPIB3: a software to solve 3D incompressible Navier–Stokes equations."
- ²¹More memory can even be spared by using an anisotropic refinement, i.e., the mesh can be refined only in the normal direction to the solid surface, while keeping a coarser mesh in the two directions parallel to the solid. It is possible that a boundary element method could avoid such difficulties, but the authors have no experience about it.
- ²²B. Lucquin and O. Pironneau, *Introduction au Calcul Scientifique* (Masson, Paris, 1996); translation in English: *Introduction to Scientific Computation*.
- ²³M.-Y. Zhou and P. Sheng, "First-principles calculations of dynamic permeability in porous media," *Phys. Rev. B* **39**, 12027–12039 (1989).
- ²⁴R. S. Maier, D. M. Kroll, Y. E. Kutsovsky, H. T. Davis, and R. S. Bernard, "Simulation of flow through bead packs using the lattice Boltzmann method," *Phys. Fluids* **10**, 60–74 (1998).
- ²⁵Remember that in such a cell, there are four spheres, each bearing twelve half necks. Using classical formulas for the volume and surface of cylinders, spheres and spherical caps yields the formula.
- ²⁶J.-F. Allard, M. Henry, J. Tizianel, L. Kelders, and W. Lauriks, "Sound propagation in air-saturated random packings of beads," *J. Acoust. Soc. Am.* **104**, 2004–2007 (1998).
- ²⁷E. Charlaix, A. P. Kushnick, and J. P. Stokes, "Experimental study of dynamic permeability in porous media," *Phys. Rev. Lett.* **61**, 1595–1598 (1988).
- ²⁸D. Singer, F. Pasierb, R. Ruel, and H. Kojima, "Multiple scattering of second sound in superfluid He II-filled porous medium," *Phys. Rev. B* **30**, 2909–2912 (1984).
- ²⁹D. L. Johnson, D. L. Hemmick, and H. Kojima, "Probing porous media with first and second sound. I. Dynamic permeability," *J. Appl. Phys.* **76**, 104–114 (1994).
- ³⁰O. Umnova, K. Attenborough, and K. M. Li, "Cell model calculations of dynamic drag parameters in packings of spheres," *J. Acoust. Soc. Am.* **107**, 3113–3119 (2000).
- ³¹D. M. J. Smeulders, R. L. G. M. Eggels, and M. E. H. van Dongen, "Dynamic permeability: reformulation of theory and new experimental and numerical data," *J. Fluid Mech.* **245**, 211–227 (1992).
- ³²J. G. Berryman and S. C. Blair, "Use of digital image analysis to estimate fluid permeability of porous materials: application of two-point correlation functions," *J. Appl. Phys.* **60**, 1930–1938 (1986).

Theory and design of sound field reproduction in reverberant rooms

Terence Betlehem and Thushara D. Abhayapala^{a)}

Department of Telecommunications Engineering, Research School of Information Sciences and Engineering, Australian National University, Canberra ACT 0200, Australia

(Received 19 July 2004; revised 6 January 2005; accepted 7 January 2005)

With the recent emergence of surround sound technology, renewed interest has been shown in the problem of sound field reproduction. However, in practical acoustical environments, the performance of sound reproduction techniques are significantly degraded by reverberation. In this paper, we develop a method of sound field reproduction for reverberant environments. The key to this method is an efficient parametrization of the acoustic transfer function over a region of space. Using this parametrization, a practical method has been provided for determining the transfer function between each loudspeaker and every point in the reproduction region. Through several simulation examples, the reverberant field designs have been shown to yield a reproduction accuracy as good as conventional free-field designs, and better than multipoint least squares designs when loudspeaker numbers are limited. The successful reproduction of sound over a wide frequency range has also been demonstrated. This approach reveals the appropriate choices for fundamental design parameters. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1863032]

PACS numbers: 43.60.Dh, 43.38.Md, 43.55.Br [NX]

Pages: 2100–2111

I. INTRODUCTION

A problem relevant to emerging surround sound technology is a reproduction of a sound field over a region of space. Using a set of loudspeakers, it is possible for listeners to spatialize sound and fully experience what it is actually like to be in the original sound environment. Sound field reproduction has been discussed since the 1960s. However, much of the work so far does not directly address sound field reproduction in reverberant environments. In this paper, using an efficient parametrization of the room transfer function we extend sound field reproduction to reverberant enclosures.

Early work in sound field reproduction was performed by Gerzon.¹ With his ambisonics system, Gerzon reproduced the first-order spherical harmonics terms of a plane wave sound field around a point in space. Ambisonics has since been unified with holography,^{2,3} both of which rely upon the Kirchoff–Helmholtz equation. Here, sound field reproduction inside a control region is achieved by controlling the pressure and its normal derivative over the boundary of the control region.⁴ In similar work, global sound field reproduction techniques^{5,6} have been proposed that control sound pressure over the boundary. By controlling sound at additional points inside the control region, these techniques obviate the need for velocity microphones. Unfortunately, such techniques require a large number of loudspeakers. For a lesser numbers of loudspeakers, Least squares techniques have been suggested by Kirkeby *et al.*^{7,8} Recently, using spherical harmonic analysis, the theoretical minimum number of loudspeakers required for the accurate reproduction of a plane wave has been established.⁹

The reverberant case is made difficult by the rapid variation of the acoustic transfer functions over the room.¹⁰ The

standard approach has been to equalize the transfer functions over multiple points using least squares techniques.^{11,12} However, in such a case, equalization can be poor away from the design points. In contrast, sound field reproduction would require the equalization to extend over the whole control region.

Alternatively, the acoustic transfer functions can be measured and incorporated into the sound field reproduction algorithm directly. Methods for estimating the acoustic transfer functions over a region have been established by Mourjopoulos¹³ and Bharitkar and Kyriakakis,¹⁴ which sample the field at a number of points and use a spatial equalization library. However, for sound field reproduction in a reverberant room, these techniques do not determine transfer functions with sufficient accuracy.

In this paper, we present a method of performing sound field reproduction in a reverberant room. This method is based on an efficient parametrization of the acoustic transfer function in the control region, where the acoustic transfer function is written as the weighted sum of a small number of orthogonal basis functions. Using this parametrization, we reconstruct a sound field accurately over the whole control region. This approach exploits the standing wave structure of the reverberant field generated by each loudspeaker to reproduce the desired sound field. We also describe a practical method for determining the acoustic transfer function between each loudspeaker and the control region, by sampling sound pressure at a small number of points.

This paper is structured as follows. In Sec. II we cast sound field reproduction into a least squares framework and introduce the basis function approach to gain insight into the fundamental parameters of the problem. In Sec. III, we describe a method of measuring the acoustic transfer functions from each speaker to any point within the control region, through measurement of the transfer function parameters of

^{a)}Electronic mail: Thushara.Abhayapala@anu.edu.au

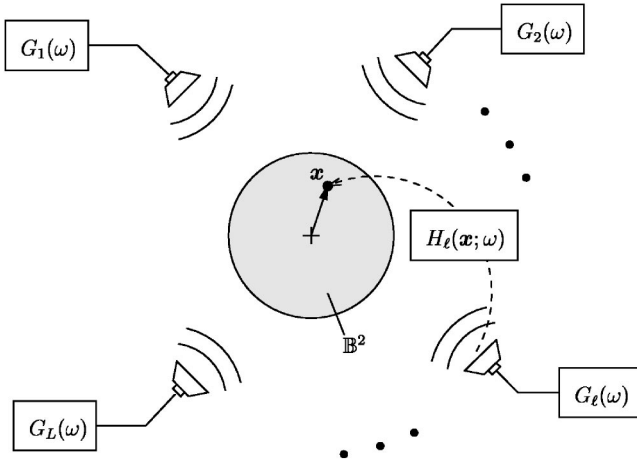


FIG. 1. Use of L loudspeakers to reproduce a desired field in a control region \mathbb{B}^2 with loudspeaker filters $G_l(\omega)$ and acoustical transfer functions $H_l(\mathbf{x}; \omega)$ from the l th loudspeaker to a point $\mathbf{x} \in \mathbb{B}^2$.

the acoustic transfer functions. Separate methods have been derived for the narrow band and wide band cases. The effect of noisy pressure samples on measurement of the parameters has also been analyzed. Finally, in Sec. IV we demonstrate the performance of our sound field reproduction technique with several examples, including a comparison with the multiple point least squares technique.

II. SOUND FIELD REPRODUCTION

In this section, we devise a method of performing two-dimensional 2-D sound field reproduction within a reverberant enclosure. This 2-D technique ensures good reproduction in the plane of the loudspeakers, provided each loudspeaker possesses a sizable vertical dimension. It is applicable to enclosures with highly sound-absorbing floors and ceilings.

The theory we develop in this paper is readily extended to 3-D space. The 2-D basis functions that are described below need only be replaced with 3-D basis functions. Unfortunately, in the case 3-D of reproduction over a volume, much larger numbers of speakers are required.⁹ We focus on reproduction in the plane as it is more practical.

Below we formulate the problem in the frequency domain. The objective is to determine the loudspeaker filter weights required to reproduce a desired sound field in a reverberant room.

A. Problem definition

We aim to reproduce the pressure $P_d(\mathbf{x}; \omega)$ of a desired sound field at each point \mathbf{x} and angular frequency ω in the source-free region of interest \mathbb{B}^2 using an array of L loudspeakers. The desired sound field could be a plane wave, a field resulting from a monopole, a field measured in a real-life scenario or the field of a surround sound system. For purposes of simplifying the analysis in this paper, we choose the control region \mathbb{B}^2 to be the circle of radius R centered about the origin:

$$\mathbb{B}^2 = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| \leq R\},$$

where $\|\cdot\|$ denotes the Euclidean distance.

As shown in Fig. 1, each loudspeaker l transmits an

output signal $G_l(\omega)$. This signal encapsulates both the input signal applied to loudspeaker l as well as any filtering of it. To characterize the acoustic properties of the enclosure, define the acoustic transfer function $H_l(\mathbf{x}; \omega)$ as the frequency response between loudspeaker l and point \mathbf{x} . $H_l(\mathbf{x}; \omega)$ summarizes the effect of reverberant reflections from the surface of the enclosure on any sound signal transmitted by each loudspeaker. The sound pressure at any point \mathbf{x} due to loudspeaker l is equal to

$$P_l(\mathbf{x}; \omega) = G_l(\omega)H_l(\mathbf{x}; \omega). \quad (1)$$

From Fig. 1, the sound pressure in the reproduced field resulting from the L loudspeakers is then equal to

$$P(\mathbf{x}; \omega) = \sum_{l=1}^L P_l(\mathbf{x}; \omega) = \sum_{l=1}^L G_l(\omega)H_l(\mathbf{x}; \omega). \quad (2)$$

The design task of sound field reproduction is to choose filter weights $G_l(\omega)$ to minimize the normalized reproduction error \mathcal{T} over \mathbb{B}^2 ,

$$\mathcal{T} = \frac{1}{\mathcal{E}} \int_{\mathbb{B}^2} |P(\mathbf{x}; \omega) - P_d(\mathbf{x}; \omega)|^2 da(\mathbf{x}), \quad (3)$$

where the normalizing factor \mathcal{E} is the energy of the desired sound field over \mathbb{B}^2 :

$$\mathcal{E} = \int_{\mathbb{B}^2} |P_d(\mathbf{x}; \omega)|^2 da(\mathbf{x}), \quad (4)$$

$da(\mathbf{x}) = x dx d\phi_x$ is the differential area element at \mathbf{x} , $x = \|\mathbf{x}\|$, and ϕ_x is the polar angle of \mathbf{x} .

The popular approach to solving this problem is to write the least squares solution over a set of uniformly spaced points over \mathbb{B}^2 . (Refs. 7, 8). A better approach is to perform the design over the whole region. This approach is proposed by Asano and Swason¹⁵ for the related problem of equalization. Yet by discretizing, these authors end up implementing a multipoint method. Below we outline a model-based approach, which uses an efficient parametrization for acoustic transfer functions to perform the design over the whole region. This model-based approach is more general than the approach of Asano and Swason¹⁵ and Santillan,¹⁶ which assume that the room is of a rectangular shape. More insight is gained into the design requirements for an accurate reproduction through the model-based approach than through multipoint least squares techniques.

B. Model-based approach

In the model-based approach, we express the sound pressure variables $P_d(\mathbf{x}; \omega)$, $P(\mathbf{x}; \omega)$, and the acoustic transfer functions $H_l(\mathbf{x}; \omega)$ in terms of the basis functions of the sound field. Provided all sound sources (including image sources produced by reflection) lie outside of \mathbb{B}^2 , at any point inside \mathbb{B}^2 the above variables can be written as a weighted sum of the inward-propagating solutions to the wave equation.¹⁷ We write the desired sound pressure $P_d(\mathbf{x}; \omega)$ as

$$P_d(\mathbf{x}; \omega) = \sum_{n=-\infty}^{\infty} \beta_n^{(d)}(\omega) J_n(kx) e^{in\phi_x}, \quad (5)$$

where $\beta_n^{(d)}(\omega)$ is the n th-order sound field coefficient of the desired sound field, $J_n(\cdot)$ is the Bessel function of order n and $k = \omega/c = 2\pi/\lambda$ is the wave number, c is the speed of sound in air, and λ is the acoustic wavelength. The functions $\{J_n(kx)e^{in\phi_x}\}_{n \in \mathbb{Z}}$ are called the basis functions of the sound field. An appropriate choice of sound field coefficients generate any valid sound field inside \mathbb{B}^2 .

Similarly, we write the reproduced sound pressure $P(\mathbf{x}; \omega)$ as

$$P(\mathbf{x}; \omega) = \sum_{n=-\infty}^{\infty} \beta_n(\omega) J_n(kx) e^{in\phi_x}, \quad (6)$$

where $\beta_n(\omega)$ are the coefficients of the reproduced sound field. A reproduction of the sound pressure $P_d(\mathbf{x}; \omega)$ over \mathbb{B}^2 with $P(\mathbf{x}; \omega)$ is equivalent to a reproduction of the sound field coefficients $\{\beta_n^{(d)}(\omega)\}_{n \in \mathbb{Z}}$ with $\{\beta_n(\omega)\}_{n \in \mathbb{Z}}$.

Because the room response is equal to the sound field pressure generated by the unit input signal $G_l(\omega)$, we can also write it as

$$H_l(\mathbf{x}; \omega) = \sum_{n=-\infty}^{\infty} \alpha_n(l, \omega) J_n(kx) e^{in\phi_x}, \quad (7)$$

where $\alpha_n(l, \omega)$ are the sound field coefficients of the room responses for loudspeaker l . These sound field coefficients completely characterize the reverberant sound field generated by each loudspeaker within \mathbb{B}^2 :

Observation 1: When the sound field coefficients $\alpha_n(l, \omega)$ for each loudspeaker are known for a given room, the acoustic transfer function $H_l(\mathbf{x}; \omega)$ between each loudspeaker and any position \mathbf{x} inside \mathbb{B}^2 is also known, and is given by Eq. (7).

Substituting Eq. (5) and Eq. (7) directly into Eq. (2), the coefficients of the reproduced sound field are related to $\alpha_n(l, \omega)$ through

$$\beta_n(\omega) = \sum_{l=1}^L G_l(\omega) \alpha_n(l, \omega). \quad (8)$$

The sequences of coefficients $[\beta_n^{(d)}(\omega)]_n$, $[\beta_n(\omega)]_n$ and $[\alpha_n(l, \omega)]$ associated with any wave field in a source-free region are shown to be bounded.¹⁸ (These coefficients are bounded in the following sense. Any field in the source-free region can be represented as the superposition of plane waves. The coefficients are bounded by the sum of the magnitudes of these plane waves.)

A benefit of the model-based approach is the ability to express key variables in terms of orthogonal functions. Using the orthogonality property of exponential functions,

$$\int_0^{2\pi} e^{-in\phi} e^{im\phi} d\phi = 2\pi \delta_{nm}, \quad (9)$$

we derive an expression for the energy \mathcal{E} and normalized error \mathcal{T} of the reproduced sound field over \mathbb{B}^2 as a function of the sound field coefficients. Starting with the field energy, we substitute Eq. (5) into Eq. (4) to yield

$$\mathcal{E} = \int_{\mathbb{B}^2} \left| \sum_{n=-\infty}^{\infty} \beta_n^{(d)}(\omega) J_n(kx) e^{in\phi_x} \right|^2 da(\mathbf{x}).$$

It follows that

$$\begin{aligned} \mathcal{E} &= \int_{\mathbb{B}^2} \sum_{n_1=-\infty}^{\infty} [\beta_{n_1}^{(d)}(\omega)]^* J_{n_1}(kx) e^{-in_1\phi_x} \\ &\quad \times \sum_{n_2=-\infty}^{\infty} \beta_{n_2}^{(d)}(\omega) J_{n_2}(kx) e^{in_2\phi_x} da(\mathbf{x}) \\ &= \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} [\beta_{n_1}^{(d)}(\omega)]^* \beta_{n_2}^{(d)}(\omega) \\ &\quad \times \int_0^{2\pi} e^{-in_1\phi_x} e^{in_2\phi_x} d\phi_x \\ &\quad \times \int_0^R J_{n_1}(kx) J_{n_2}(kx) x dx, \end{aligned}$$

where we have applied $da(\mathbf{x}) = x dx d\phi_x$ in the second step. Applying the orthogonality property, Eq. (9), of the exponential functions, the field energy reduces to

$$\begin{aligned} \mathcal{E} &= 2\pi \sum_{n=-\infty}^{\infty} |\beta_n^{(d)}(\omega)|^2 \int_0^R [J_n(kx)]^2 x dx \\ &= K \sum_{n=-\infty}^{\infty} w_n(kR) |\beta_n^{(d)}(\omega)|^2, \end{aligned} \quad (10)$$

where $K = 2\pi/k^2$ and

$$w_n(kR) \triangleq k^2 \int_0^R [J_n(kx)]^2 x dx = \int_0^{kR} [J_n(x)]^2 x dx. \quad (11)$$

The second step was performed with the variable substitution $x' = kx$. Similarly, substituting Eq. (5) and Eq. (6) into Eq. (3), the normalized error becomes

$$\mathcal{T} = \frac{1}{\mathcal{E}} \int_{\mathbb{B}^2} \left| \sum_{n=-\infty}^{\infty} [\beta_n^{(d)}(\omega) - \beta_n(\omega)] J_n(kx) e^{in\phi_x} \right|^2 da(\mathbf{x}).$$

Utilizing the orthogonality property, the normalized error reduces to

$$\mathcal{T} = \frac{K}{\mathcal{E}} \sum_{n=-\infty}^{\infty} w_n(kR) |\beta_n^{(d)}(\omega) - \beta_n(\omega)|^2. \quad (12)$$

We shall call $w_n(kR)$ in (12) the *coefficient weighting function*.

Since the summations in Eq. (5), Eq. (6), and Eq. (7) have infinite numbers of terms, it may seem that the above parametrizations need an infinite number of coefficients. However, in the next section, we show that for any finite control region, they need only a finite number of coefficients to accurately represent a sound field or a room response.

C. Active basis functions

Because of the high-pass character of Bessel functions, not all of the basis functions make a significant energy contribution to the sound field inside \mathbb{B}^2 . Studying Eq. (12), because the sequences of sound field coefficients $[\beta_n^{(d)}(\omega)]_n$ and $[\beta_n(\omega)]_n$ are bounded while from Fig. 2,

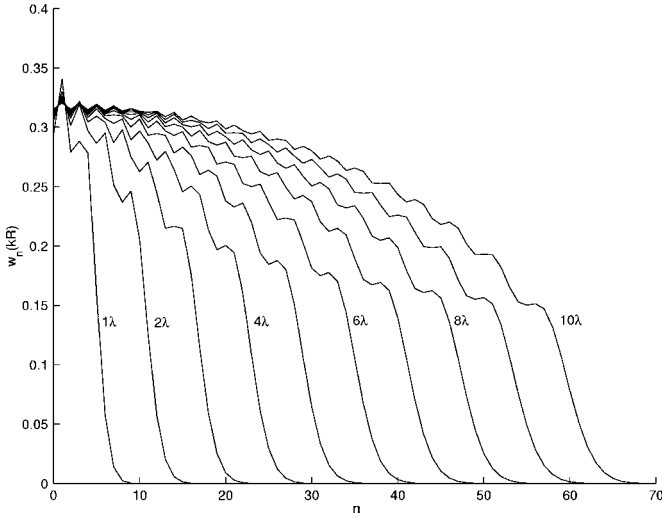


FIG. 2. A plot of the coefficient weighting function $w_{|n|}(kR)$ vs $|n|$ for control regions with radii $R=[1\lambda, 2\lambda, \dots, 10\lambda]$. λ is the acoustic wavelength of interest, related to the wave number by $k=2\pi/\lambda$.

$w_n(kR)$ drop rapidly to zero past a threshold, the energy contribution of each term to reproduction error is controlled by $w_n(kR)$.

Previous work^{9,18} asserts that only the basis functions of index up to $N=[kR]$ contribute significant energy to the sound field inside B^2 . This result is supported by Fig. 2, where $w_{|n|}(kR)$ is plotted against $|n|$. [Note the coefficient weighting functions of negative index are a mirror of those of positive index $w_{-n}(kR)=w_n(kR)$. This can be seen by applying the Bessel function property $J_n(x)=(-1)^n J_{-n}(x)$.] The weighting is seen to be small past $|n|>N$. The $2N+1$ basis functions, $J_{-N}(kx)e^{-iN\phi_x}, \dots, J_N(kx)e^{iN\phi_x}$ are referred to as *active* in B^2 . The remaining basis functions are referred to as *inactive* in B^2 . Such basis functions only contribute significant energy to the sound field outside of B^2 .

An accurate sound field reproduction requires the reproduction of these active basis functions. Also, the acoustic transfer functions mentioned in Observation 1 are accurately determined just by measuring the sound field coefficients $\{\alpha_n(l, \omega)\}_{n=-N}^N$ of the active basis functions.

D. Least squares solution

We now derive the least squares solution for the speaker filter weights that minimizes the reproduction error in Eq. (12). This solution is expressed in terms of the sound field coefficients.

Because the weighting of terms in the normalized error in Eq. (12) rapidly diminish for $|n| \geq N$, it can be truncated to

$$\mathcal{T}_{N_T} = \frac{K}{\mathcal{E}} \sum_{n=-N_T}^{N_T} w_n(kR) |\beta_n(\omega) - \beta_n^{(d)}(\omega)|^2, \quad (13)$$

for $N_T \geq N$. This truncated reproduction error \mathcal{T}_{N_T} can be written in matrix form, as follows. Defining the vector of loudspeaker filter weights $\mathbf{g}=[G_1(\omega), G_2(\omega), \dots, G_L(\omega)]^T$, where $[\cdot]^T$ is the matrix transpose operator, the vector of the coefficients of the reproduced sound field $\boldsymbol{\beta}$

$=[\beta_{-N_T}(\omega), \beta_{-N_T+1}(\omega), \dots, \beta_{N_T}(\omega)]^T$, and the matrix of the coefficients of the room responses of all loudspeakers,

$$\mathbf{A} = \begin{bmatrix} \alpha_{-N_T}(1, \omega) & \alpha_{-N_T}(2, \omega) & \cdots & \alpha_{-N_T}(L, \omega) \\ \alpha_{1-N_T}(1, \omega) & \alpha_{1-N_T}(2, \omega) & \cdots & \alpha_{1-N_T}(L, \omega) \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{N_T}(1, \omega) & \alpha_{N_T}(2, \omega) & \cdots & \alpha_{N_T}(L, \omega) \end{bmatrix}, \quad (14)$$

Eq. (8) can be rewritten as $\boldsymbol{\beta}=\mathbf{A}\mathbf{g}$. Additionally, define the vectors of the coefficients of the desired sound field, $\boldsymbol{\beta}_d=[\beta_{-N_T}^{(d)}(\omega), \beta_{-N_T+1}^{(d)}(\omega), \dots, \beta_{N_T}^{(d)}(\omega)]^T$, and the diagonal weighting matrix,

$$\mathbf{W} = \begin{bmatrix} w_{-N_T}(kR) & 0 & \cdots & 0 \\ 0 & w_{-N_T+1}(kR) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{N_T}(kR) \end{bmatrix}.$$

Writing the numerator of Eq. (13) in matrix form:

$$\sum_{n=-N_T}^{N_T} w_n(kR) |\beta_n(\omega) - \beta_n^{(d)}(\omega)|^2 = (\boldsymbol{\beta} - \boldsymbol{\beta}_d)^H \mathbf{W} (\boldsymbol{\beta} - \boldsymbol{\beta}_d),$$

where $(\cdot)^H$ is the matrix Hermitian operator, the truncated reproduction error can be written as

$$\mathcal{T}_{N_T} = \frac{(\boldsymbol{\beta} - \boldsymbol{\beta}_d)^H \mathbf{W} (\boldsymbol{\beta} - \boldsymbol{\beta}_d)}{\boldsymbol{\beta}_d^H \mathbf{W} \boldsymbol{\beta}_d}.$$

Since $\boldsymbol{\beta}=\mathbf{A}\mathbf{g}$, we expand the truncated reproduction error as a quadratic form in the vector of loudspeaker filter weights:

$$\mathcal{T}_{N_T}(\mathbf{g}) = \frac{1}{d} (\mathbf{g}^H \mathbf{B} \mathbf{g} - \mathbf{b}^H \mathbf{g} - \mathbf{g}^H \mathbf{b} + d),$$

where $\mathbf{B}=\mathbf{A}^H \mathbf{W} \mathbf{A}$, $\mathbf{b}=\mathbf{A}^H \mathbf{W} \boldsymbol{\beta}_d$, $d=\boldsymbol{\beta}_d^H \mathbf{W} \boldsymbol{\beta}_d$. This quadratic form possesses its global minimum at¹⁵

$$\hat{\mathbf{g}} = \mathbf{B}^{-1} \mathbf{b} = (\mathbf{A}^H \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{W} \boldsymbol{\beta}_d, \quad (15)$$

with the associated minimum in truncated reproduction error:

$$\mathcal{T}_{N_T}(\hat{\mathbf{g}}) = 1 - \frac{1}{d} \mathbf{b}^H \mathbf{B}^{-1} \mathbf{b}.$$

Once $(\mathbf{A}^H \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{W}$ is computed for the acoustical environment, the reproduced sound field can be changed easily by modifying $\boldsymbol{\beta}_d$.

E. Multiple-point approach

For comparison, we describe the conventional least squares approach, where the sound field is reproduced at several points. Here we aim to reproduce the desired sound field $P_d(\mathbf{x}; \omega)$ over M points $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M$, positioned within the region of interest B^2 with $M \geq L$. Define the vector of desired sound pressure at each point $\mathbf{p}_d=[P_d(\mathbf{x}_1; \omega), P_d(\mathbf{x}_2; \omega), \dots, P_d(\mathbf{x}_M; \omega)]^T$, and the acoustic transfer functions between the loudspeakers and control points into the matrix,

$$\mathbf{H} = \begin{bmatrix} H_1(\mathbf{x}_1; \omega) & H_2(\mathbf{x}_1; \omega) & \cdots & H_L(\mathbf{x}_1; \omega) \\ H_1(\mathbf{x}_2; \omega) & H_2(\mathbf{x}_2; \omega) & \cdots & H_L(\mathbf{x}_2; \omega) \\ \vdots & \vdots & \ddots & \vdots \\ H_1(\mathbf{x}_M; \omega) & H_2(\mathbf{x}_M; \omega) & \cdots & H_L(\mathbf{x}_M; \omega) \end{bmatrix}.$$

The loudspeaker weights are determined from the (possibly overdetermined) system $\mathbf{H}\mathbf{g}=\mathbf{p}$. The least squares solution is then given by $\hat{\mathbf{g}}=\mathbf{H}^\dagger\mathbf{p}$, where $[\cdot]^\dagger$ is the Moore–Penrose inverse.

In Sec. IV, the multiple-point approach has been compared to the model-based approach. We shall see that since the model-based approach targets reproduction over the whole control region, it yields superior performance to the multiple-point approach.

In the next section we describe a method for measuring coefficients for the acoustic transfer function in matrix \mathbf{A} for the model-based approach.

III. ESTIMATION OF SOUND FIELD COEFFICIENTS

In this section we describe how to fully determine the sound field inside a control region \mathbb{B}^2 through measurement of the sound field coefficients. This task is important as it is required to calculate $\{\alpha_n(l, \omega)\}_{n \in \mathbb{Z}}$ that characterize the reverberant field generated by each loudspeaker.

We write the sound pressure $P(\mathbf{x}; \omega)$ inside \mathbb{B}^2 generated by a loudspeaker outside \mathbb{B}^2 in a reverberant enclosure as the basis function expansion:

$$P(\mathbf{x}; \omega) = \sum_{n=-\infty}^{\infty} \beta_n(\omega) J_n(kx) e^{in\phi_x}, \quad (16)$$

where $\beta_n(\omega)$ is the sound field coefficient of order n . To determine the field pressure inside \mathbb{B}^2 , we describe a simple means of measuring $\beta_n(\omega)$.

The method used to determine the sound field coefficients varies depending on whether they are required in a narrow range of frequencies (Sec. III A) or a wide range of frequencies (Sec. III B).

A. Narrow-band method

In the case that sound field reproduction is performed in a narrow frequency range for a choice of R , away from any zero of $J_n(kR)$, good sound field coefficient estimates are obtained by sampling pressure over a single circle of radius R .

1. Computation of sound field coefficients

The sound field coefficients are obtained from the analysis equation,

$$\beta_n(\omega) = \frac{1}{2\pi J_n(kx)} \int_0^{2\pi} P(\mathbf{x}; \omega) e^{-in\phi_x} d\phi_x, \quad (17)$$

provided x is not a zero of $J_n(kx)$. This equation is derived by multiplying both sides of Eq. (16) by $e^{-in'\phi}$, integrating, and applying the orthogonality property. Interpreting this equation, the sound field coefficients and hence sound field can be known in the whole of \mathbb{B}^2 just by measuring sound pressure on a circle of radius x .

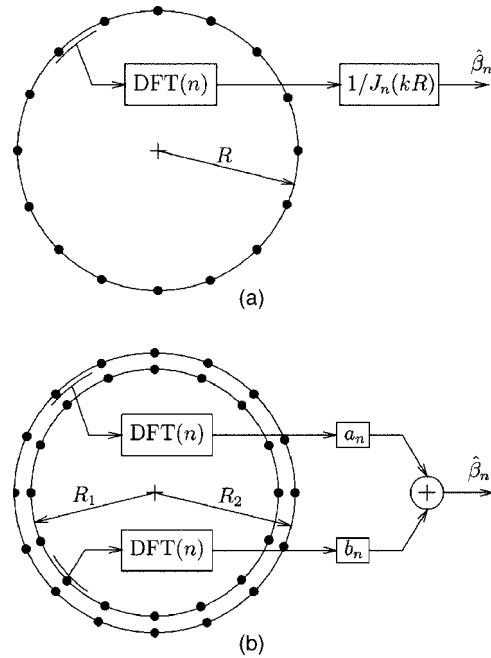


FIG. 3. Proposed methods for measuring the sound field coefficients $\hat{\beta}_n(\omega)$ in (a) the narrow band case where pressure is sampled at one radius R and (b) the wide band case where pressure is sampled at two radii, R_1 and R_2 .

In this paper we sample pressure at $x=R$, on the boundary of \mathbb{B}^2 . Now at a radius x , only basis functions of order up to $[kx]$ are active. Over the boundary all of the active basis functions of \mathbb{B}^2 are active, while the higher-order basis functions are inactive. So heuristically this choice of sampling radius makes sense.

Approximate sound field coefficients $\hat{\beta}_n(\omega)$ are obtained by sampling sound pressure at M evenly spaced points (R, ϕ_m) , where $\phi_m = 2\pi m/M$ for $m=0, 1, \dots, M-1$. Since Eq. (17) shows $\beta_n(\omega) J_n(kR)$ are the Fourier series coefficients of $P(R, \phi; \omega)$ in variable ϕ . Consequently, it can be approximated with the discrete Fourier transform (DFT) relationship:

$$\hat{\beta}_n(\omega) = \frac{1}{J_n(kR)} \text{DFT}\{P(R, \phi_m; \omega)\}(n), \quad (18)$$

where $\text{DFT}\{f(m)\}(n)$ is the M -point DFT, defined by

$$\text{DFT}\{f(m)\}(n) = \frac{1}{M} \sum_{m=0}^{M-1} f(m) e^{-i(2\pi mn/M)}. \quad (19)$$

Coefficients $\hat{\beta}_n(\omega)$ are recognized as the DFT of the sampled field pressure around the circle $\{P(R, 2\pi m/M; \omega)\}_{m=0}^{M-1}$, weighted by the Bessel function term $1/J_n(kR)$ [Fig. 3(a)].

An appropriate choice for M can be deduced by noting that the sound field in \mathbb{B}^2 is the result of $2N+1$ active basis functions. Since one equation is required for each sound field coefficient, we need at least $M=2N+1$ pressure samples where $N=\lceil kR \rceil$.

Due to the presence of $1/J_n(kR)$ in Eq. (18), if kR is near one of the Bessel zeros, coefficient error is amplified. This error amplification can be negated by oversampling pressure, as is seen in the next section.

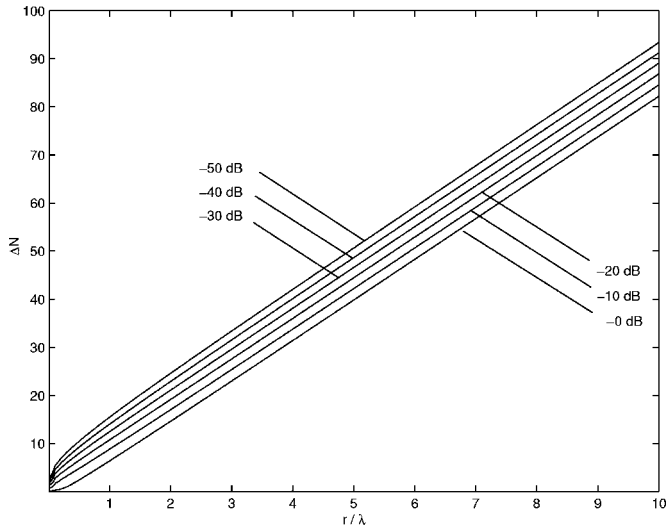


FIG. 4. Sample measurement parameter ΔN required for several values of ϵ' (in dB). The curves show the ΔN required to ensure that $|J_{\Delta N}(kR)|$ is upper bounded by ϵ' . The total number of pressure samples required is then $N + [\Delta N]$, where N is the number of active basis functions.

2. Approximation error

In Appendix A, we show the error in the approximated sound field coefficients $\hat{\beta}_n(\omega)$ is given by

$$\hat{\beta}_n(\omega) - \beta_n(\omega) = \frac{1}{J_n(kR)} \sum_{q=-\infty, \neq 0}^{\infty} J_{n+qM}(kR) \beta_{n+qM}(\omega). \quad (20)$$

This equation reveals a type of aliasing, since the higher-order coefficients $\{\beta_{n+qM}(\omega)\}_{q=-\infty, \neq 0}^{\infty}$ are mapped onto each $\hat{\beta}_n(\omega)$. It also shows the magnitude of the approximation error is controlled by the size of $1/J_n(kR)$. We shall refer to the summation term in Eq. (20) as the *aliasing error* and the preceding $1/J_n(kR)$ term as the *error scaling*. If $J_n(kR)$ is small, the error scaling is large.

From Eq. (20), each basis function of index $n+qM$ makes a termwise contribution of $[J_{n+qM}(kR)/J_n(kR)]\beta_{n+qM}(\omega)$ for $q \neq 0$ to the approximation error in $\hat{\beta}_n(\omega)$. We now identify a choice of M that ensures the termwise scale factor $J_{n+qM}(kR)/J_n(kR)$ arbitrarily small.

Define κ_n as the largest *termwise scale factor* in $\beta_n(\omega)$:

$$\kappa_n \triangleq \max_{q=-\infty, \dots, \infty, \neq 0} \left| \frac{J_{n+qM}(kR)}{J_n(kR)} \right|, \quad (21)$$

for $n = -N, -N+1, \dots, N$. In Appendix B, the largest termwise scale factor of all of the active basis functions is shown to be bounded by

$$\begin{aligned} \max_{n=-N, \dots, N} \kappa_n &\leq \underbrace{\max_{n=0, \dots, N} 1/|J_n(kR)|}_{\kappa_{\text{es}}} \\ &\times \underbrace{\frac{1}{\sqrt{2\pi(M-N)}} \left[\frac{ekR}{2(M-N)} \right]^{M-N}}_{\text{upper bound on } |J_{M-N}(kR)|}. \end{aligned} \quad (22)$$

The first term in this bound is the maximum error scaling that we shall denote as κ_{es} . The second term is a bound on the Bessel function $J_{M-N}(kR)$ obtained from Jones *et al.*¹⁸ We note in Eq. (22) that the largest termwise scaling factor decays exponentially to zero as M is increased past $N + [ekR/2]$. This observation suggests choosing $M \approx N + [ekR/2]$. However, a better procedure for the choice of M , motivated by the form of Eq. (22) is presented next.

3. Conservative estimate of M

This procedure allows a more accurate choice of M .

- Choose the desired bound ϵ on the termwise scale factor; i.e., choose a bound for which $\max_{n=-N, \dots, N} \kappa_n < \epsilon$.
- Calculate the maximum error scaling:

$$\kappa_{\text{es}} = \max_{n=0, 1, \dots, N} |1/J_n(kR)|.$$
- Determine $\Delta N = M - N$ to guarantee $|J_{N-M}(kR)|$ is upper bounded by $\epsilon' = \epsilon/\kappa_{\text{es}}$ through the relationship

$$\frac{1}{\sqrt{2\pi\Delta N}} \left[\frac{ekR}{2\Delta N} \right]^{\Delta N} = \epsilon'. \quad (23)$$
- The required number of samples $M = N + [\Delta N]$.

A judicious choice of radius R ensures that κ_{es} is minimal. Further, such a choice will result in minimizing the number of required pressure samples.

Interestingly, Fig. 4 shows a linear relationship between ΔN and kR for large kR . Rearranging Eq. (23):

$$\Delta N = \frac{e}{2} (\sqrt{2\pi\Delta N\epsilon'})^{-1/\Delta N} kR.$$

As $\Delta N \rightarrow \infty$, the term $(\sqrt{2\pi\Delta N\epsilon'})^{-1/\Delta N} \rightarrow 1$, causing this expression to reduce to $\Delta N \approx ekR/2$. This relationship explains the linear section of the curves in Fig. 4 and is consistent with the $N + [ekR/2]$ rule.

In summary, we require at least $M = 2N + 1$ pressure samples to measure the sound field coefficients of the active basis functions. An analysis of the error in approximated sound field coefficients shows that for an accurate measurement of sound field coefficients more pressure samples may be required. The larger M is required to negate the effects of error scaling. A conservative procedure for estimating M is summarized above.

B. Wide band method

In frequency ranges and sizes of \mathbb{B}^2 of interest to practical problems, the Bessel term $J_n(kR)$ is guaranteed to be zero at a number of frequencies. These zeros cause problems when designing with the above method over a wide frequency range. For each zero, a basis functions remains unmeasurable.

To illustrate the magnitude of the problem, consider the asymptotic behavior (that is, the behavior for large kR) of the Bessel function:¹⁷

$$J_n(kR) \sim \sqrt{\frac{2}{\pi kR}} \cos(kR - n\pi/2 - \pi/4). \quad (24)$$

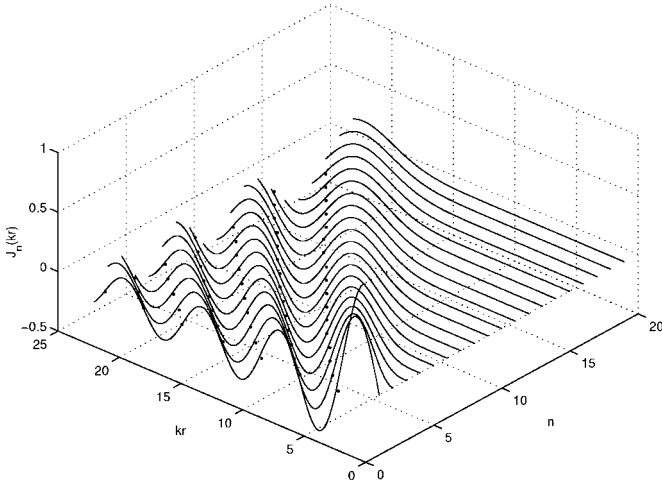


FIG. 5. Bessel functions $J_n(kR)$ that are active in a control region of radius 0.3 m and frequencies up to 4 kHz. Each zero of the Bessel functions is marked with a dot (\cdot).

One can see from this equation that each of the Bessel functions has zeros spaced approximately $kR = \pi$ or $f = c/2R$ apart. In a region B^2 with radius $R = 0.3$ m and speed of sound $c = 342$ m/s, each $J_n(kR)$ has zeros spaced 570 Hz apart. In a 0–4 kHz frequency range there are 58 zeros present (Fig. 5). On average, one zero occurs every 69 Hz, with the larger concentration of zeros at higher frequencies.

To combat this problem, we propose an alternative method for the wide-band case. Instead of sampling over a single radius R we sample over two concentric circles of radii $R_1 = R - \delta R$ and $R_2 = R$ [Fig. 3(b)].

1. Computation of sound field coefficients

Multiplying both sides of Eq. (16) by basis function $J_{n'}(kx)e^{-in'\phi_x}$ and integrating over the thin shell of thickness $\delta R = R_2 - R_1$, $\{\mathbf{x} \in \mathbb{R}^2: R_1 \leq \|\mathbf{x}\| \leq R_2\}$, the orthogonality property, Eq. (9), is used to show that

$$\beta_n(\omega) = \frac{1}{2\pi \int_{R_1}^{R_2} [J_n(kx)]^2 x dx} \times \int_{R_1}^{R_2} \int_0^{2\pi} P(\mathbf{x}; \omega) J_n(kx) e^{-in\phi_x} dx d\phi_x.$$

For small δR , we can approximate the integral in x with the zeroth-order approximation:

$$\int_{R_1}^{R_2} f(x) dx \approx \frac{1}{2} [f(R_1) + f(R_2)] \delta R. \quad (25)$$

Using Eq. (25), we express $\beta_n(\omega)$ as a sum of two weighted Fourier series equations. Sampling the field with M evenly spaced sensor pairs positioned at (R_1, ϕ_m) and (R_2, ϕ_m) , the sound field coefficients $\beta_n(\omega)$ are estimated with

$$\hat{\beta}_n(\omega) = a_n(\omega) \text{DFT}\{P(R_1, \phi_m; \omega)\}(n) + b_n(\omega) \text{DFT}\{P(R_2, \phi_m; \omega)\}(n), \quad (26)$$

where for $R_1 \approx R_2$, the DFTs are weighted by

$$a_n(\omega) = \frac{J_n(kR_1)}{[J_n(kR_1)]^2 + [J_n(kR_2)]^2}, \quad (27)$$

$$b_n(\omega) = \frac{J_n(kR_2)}{[J_n(kR_1)]^2 + [J_n(kR_2)]^2}. \quad (28)$$

The approximated sound field coefficients can hence be obtained through taking the DFT of the pressure samples around each circle and calculating a weighted average.

Next we analyze the error in the approximated sound field coefficients.

2. Approximation error

For the wide band method, the error in the approximated sound field coefficients is

$$\hat{\beta}_n(\omega) - \beta_n(\omega) = a_n(\omega) \sum_{q=-\infty, \neq 0}^{\infty} J_{n+qM}(kR_1) \beta_{n+qM}(\omega) + b_n(\omega) \sum_{q=-\infty, \neq 0}^{\infty} J_{n+qM}(kR_2) \beta_{n+qM}(\omega). \quad (29)$$

This expression is proven by substituting Eq. (19) into Eq. (26) and simplifying the resulting expression in a manner similar to that in Appendix A. In contrast to the narrow band case in Eq. (20), the wide band case possesses two error scaling terms, $a_n(\omega)$ and $b_n(\omega)$. Also, the presence of $[J_n(kR_1)]^2 + [J_n(kR_2)]^2$ in the denominators of the error scaling terms [see Eq. (27) and Eq. (28)] improves the robustness at the zeros.

The critical parameter in the wide band technique is δR . δR controls the maximum value of the error scaling terms $a_n(\omega)$ and $b_n(\omega)$, as we will now show. When either kR_1 or kR_2 is a zero of the Bessel function, approximation error simplifies to the narrow band expression in Eq. (20). In the case $J_n(kR_1) = 0$, the error scaling terms reduce to $a_n(\omega) = 0$ and $b_n(\omega) = 1/|J_n(kR_2)|$. For δR small, $J_n(kR_2)$ is also small and the linear approximation $J_n(kR_2) = k\delta R J'_n(kR_1)$ can be made. By the derivative property of the Bessel function $xJ'_n(x) = nJ_n(x) - xJ_{n+1}(x)$, we see that $J'_n(kR_1) = J_{n+1}(kR_1)$, so the nonzero error scaling term is

$$b_n(\omega) \approx \frac{1}{k\delta R |J_{n+1}(kR_1)|}.$$

From this equation, it seems advantageous to choose δR large, as a larger δR implies a smaller error scaling. However, δR cannot be too large, otherwise $J_n(kR_2)$ may coincide with another zero of the same Bessel function. As these Bessel functions are regularly spaced, we can select δR to avoid this case. From Eq. (24) the Bessel zeros of $J_n(kR)$ are spaced π apart, so set $k\delta R < \pi/2$ or $\delta R < \lambda/4$. An appropriate choice of δR is hence 1/4 of the smallest acoustic wavelength of interest.

C. Impact of measurement noise

In real rooms with background noise and sensor noise, it is nontrivial to obtain clean measurements of the acoustical

transfer functions. In this section we study how such noise impacts measurement of the sound field coefficients.

Model the measurement noise $\eta(\mathbf{x};\omega)$ at each sensor position \mathbf{x} as additive white noise of zero mean and variance $\sigma^2(\omega)$. The noisy pressure is

$$\tilde{P}(\mathbf{x};\omega) = P(\mathbf{x};\omega) + \eta(\mathbf{x};\omega). \quad (30)$$

Calculating the DFT of both sides of Eq. (30) and comparing with Eq. (18), the noisy sound field coefficient estimates $\tilde{\beta}_n(\omega)$ are shown to be related to the noiseless estimates $\hat{\beta}_n(\omega)$ by

$$\tilde{\beta}_n(\omega) = \hat{\beta}_n(\omega) + \frac{1}{J_n(kR)} \text{DFT}\{\eta(R, \phi_m; \omega)\}(n).$$

Inserting the definition of the DFT in Eq. (19) and rearranging:

$$\tilde{\beta}_n(\omega) - \hat{\beta}_n(\omega) = \frac{1}{J_n(kR)} \frac{1}{M} \sum_{m=0}^{M-1} \eta(R, \phi_m; \omega) e^{-i(2\pi mn/M)}. \quad (31)$$

Equation (31) is used to derive the mean and variance of the noisy sound field coefficient estimates. Taking the expectation of both sides of Eq. (31), the zero mean property of $\eta(R, \phi_m; \omega)$ implies that

$$E\{\tilde{\beta}_n(\omega) - \hat{\beta}_n(\omega)\} = 0,$$

or $E\{\tilde{\beta}_n(\omega)\} = \hat{\beta}_n(\omega)$. Measurement of the sound field coefficients remains unbiased by noise with zero mean. Multiplying Eq. (31) by its complex conjugate and taking the expectation, the variance is given by

$$\begin{aligned} E\{|\tilde{\beta}_n(\omega) - \hat{\beta}_n(\omega)|^2\} &= \frac{1}{[J_n(kR)]^2} \frac{1}{M^2} \\ &\times \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} E\{\eta^*(R, \phi_{m_1}; \omega) \\ &\times \eta(R, \phi_{m_2}; \omega)\} \\ &\times \exp\{i2\pi(m_1 - m_2)n/M\}. \end{aligned}$$

In the case that noise is spatially uncorrelated, $E\{\eta^*(R, \phi_{m_1}; \omega) \eta(R, \phi_{m_2}; \omega)\} = \sigma^2(\omega) \delta_{m_1 m_2}$, and the variance reduces to

$$E\{|\tilde{\beta}_n(\omega) - \hat{\beta}_n(\omega)|^2\} = \frac{1}{M} \frac{\sigma^2(\omega)}{J_n(kR)^2}.$$

The variance is influenced by error scaling factor $1/|J_n(kR)|$. In the wide band case, we can use a similar derivation to show that the sound field coefficient estimates are also unbiased and have a variance given by

$$E\{|\tilde{\beta}_n(\omega) - \hat{\beta}_n(\omega)|^2\} = \frac{1}{M} \frac{\sigma^2(\omega)}{[J_n(kR_1)]^2 + [J_n(kR_2)]^2}. \quad (32)$$

The Bessel functions in the denominators of Eq. (32) show that similar error scaling occurs in the noise error of the wide band case.

This error scaling of the measurement noise impacts the

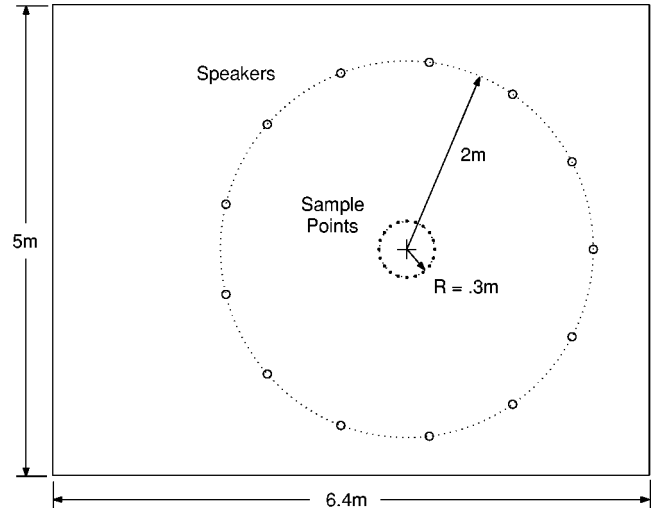


FIG. 6. Layout of loudspeakers (○) and pressure sample points (●) for sound field reproduction in the simulations. Though loudspeakers could be arbitrarily placed, we set them on a circle of radius 2 m centered about [3.8 m, 2.4 m].

measurability of the sound field coefficients. In general, when $J_n(kR)$ is small the error scaling causes a large amplification of measurement noise. For the inactive basis functions, the Bessel terms $J_n(kR)$ are so small as to be effectively zero. The resulting error scaling is so large that the sound field coefficients are unmeasurable, even for a small $\sigma^2(\omega)$.

IV. SIMULATION EXAMPLES

In the following examples, we illustrate the sound reproduction of a plane wave and a single monopole source at a single frequency (Sec. IV A and Sec. IV B) and at a range of frequencies (Sec. III D). Single frequency reproduction is performed at 1 kHz. Then in Sec. IV C we examine the performance of reproduction for the case that the numbers of loudspeakers are inadequate. In Sec. IV D we demonstrate the influence of measurement noise on reproduction error.

The reverberant room parameters and loudspeaker placement are summarized in Fig. 6. The room is rectangular with a wall absorption coefficient of 0.3. Unless otherwise stated, the control region has a radius of 0.3 m. Though the sound field reproduction design technique is applicable for any configuration and type of loudspeaker, we perform the sound field reproduction with a circular array of omnidirectional loudspeakers. This setup yields an average direct-to-reverberant energy ratio from each loudspeaker of -4.4 dB at the center of B^2 .

The loudspeaker requirements of this scenario are governed by the control region parameter $N = \lceil kR \rceil = 6$, prompting the use of $2N + 1 = 13$ loudspeakers. Following the conservative design procedure of Sec. IV A with $\epsilon = -20$ dB, the maximum error scaling is $\kappa = 25$ dB, and from Fig. 4 the $[\Delta N]$ corresponding to $\kappa' = \epsilon/\kappa$ is 14. We hence sample the pressure at $M = N + [\Delta N] = 20$ points to measure the room response coefficients $\alpha_n(l, \omega)$.

The reverberation is generated with a 2-D adaptation of

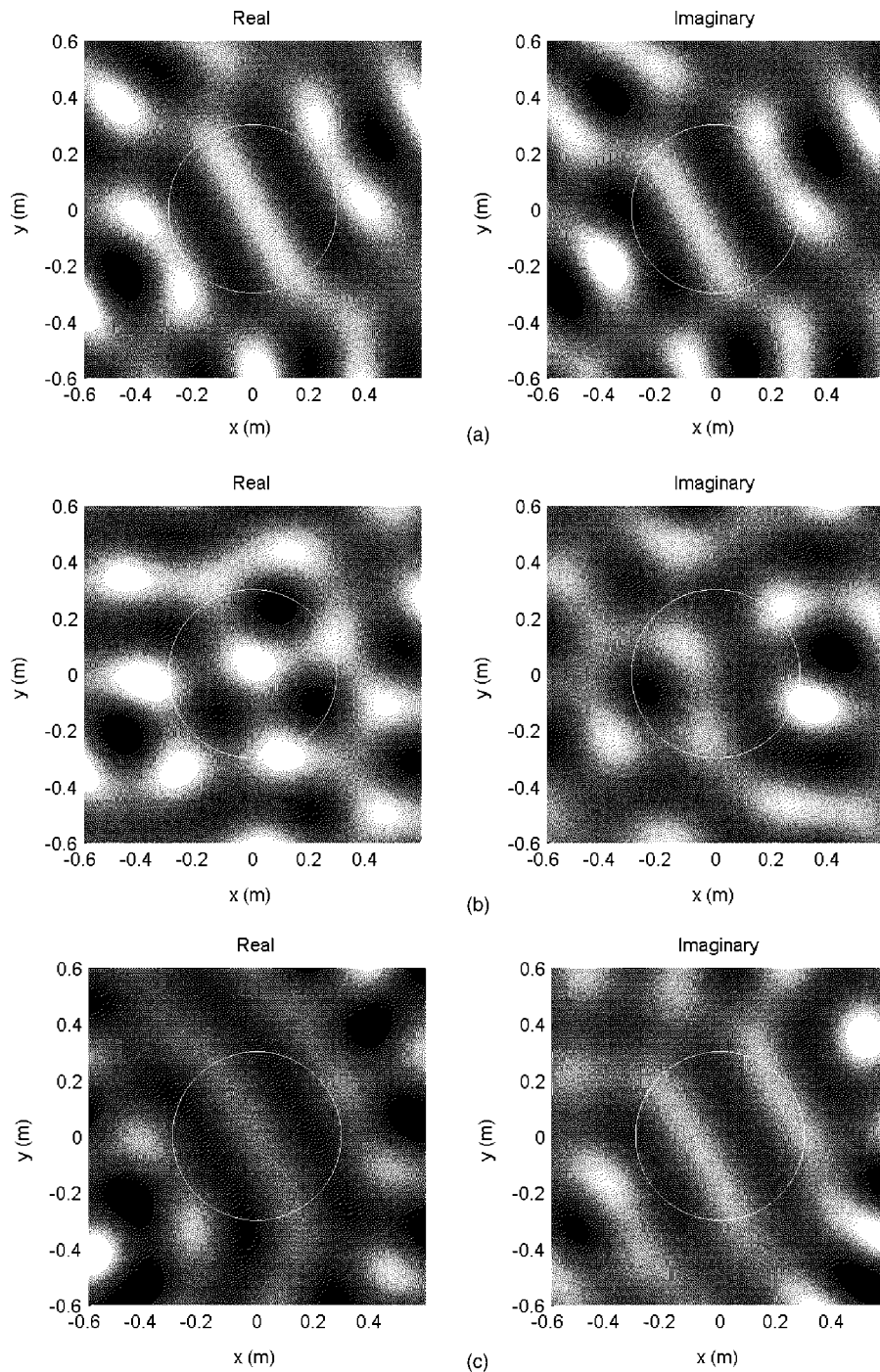


FIG. 7. The reproduction of a plane wave with 13 speakers and 20 pressure samples in a 0.3 m radius circle, for (a) a free field, (b) the same free field design in the reverberant room, and (c) a reverberant field design in the reverberant room. Reproduction errors are 0.87%, 307%, and 0.85%, respectively.

the image-source method.¹⁹ Each of the room frequency response functions is given by

$$H(\mathbf{x}; \omega) = \mathcal{H}_0^{(2)}(k\|\mathbf{x} - \mathbf{y}\|) + \sum_{n=1}^{N_i} \zeta_n \mathcal{H}_0^{(2)}(k\|\mathbf{x} - \mathbf{y}_n\|),$$

where $\mathcal{H}_0^{(2)}(\cdot)$ is the zeroth-order Hankel function of the second kind, \mathbf{y} is the source position, and ζ_n and \mathbf{y}_n are the position and accumulated reflection coefficient of the n th image-source, respectively. [For the 2-D point source, or a cylindrical source, $\mathcal{H}_0^{(2)}(kr)$ gives the field at a distance r from the source;²⁰ for a 3-D point source, this is equal to the more familiar expression $h_0^{(2)}(kr) = ie^{-ikr}/kr$.] Image-source positions are obtained through the repeated mirroring

about the walls of the enclosure.¹⁹ In simulations below, all the image-sources of up to fifth order were included (totaling $N_i = 60$ image sources).

Sound field reproduction results are illustrated in Figs. 7–9. Here the real and imaginary parts of the complex pressure of the reproduced field are displayed as density plots. Details of the sound field reproduction in each case are described below.

A. Reproduction of a plane wave

First, the field pressure of a plane wave is reproduced. For a plane wave originating from direction $\hat{\mathbf{y}}$:

$$P_d(\mathbf{x}; \omega) = e^{-ik\mathbf{x} \cdot \hat{\mathbf{y}}}.$$

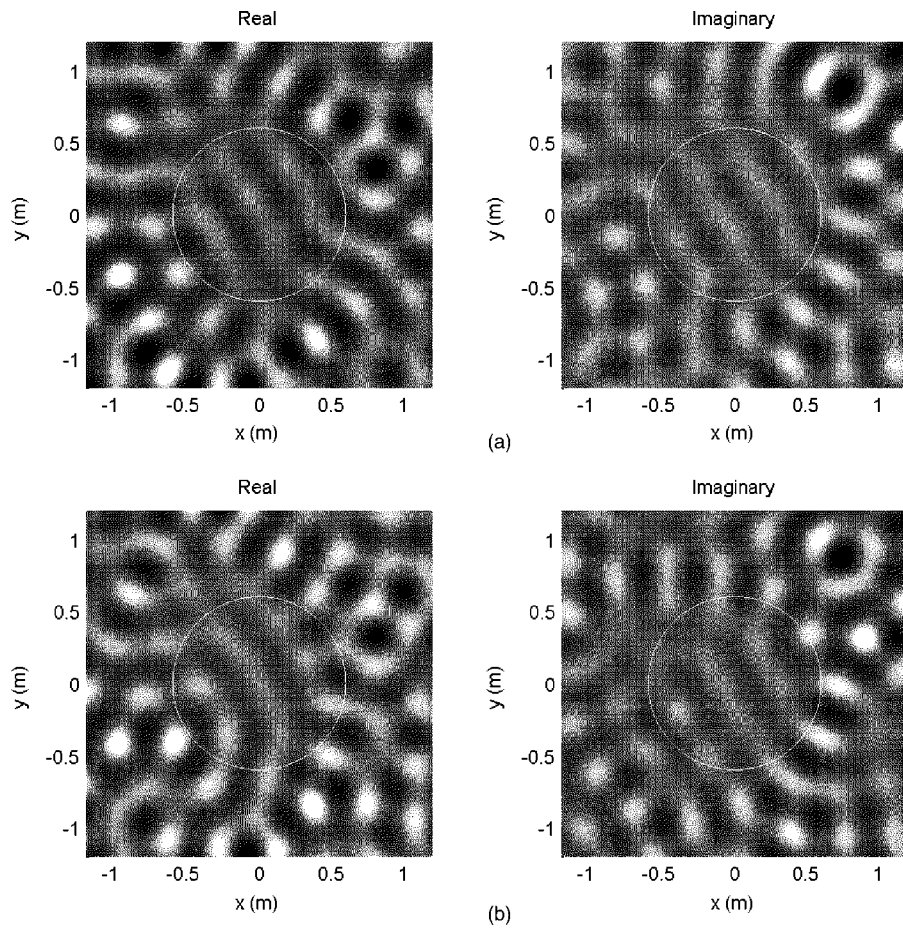


FIG. 8. The reproduction of a monopole in a 0.3 m radius circle of the reverberant room with 13 speakers and 20 pressure samples. The reproduction error is 2.12%. The position of the monopole is marked with a “+.”

Through the Jacobi–Anger expression,²⁰

$$e^{-ik\mathbf{x}\cdot\hat{\mathbf{y}}} = \sum_{n=-\infty}^{\infty} (-i)^n e^{-in\phi_y} J_n(kx) e^{in\phi_x},$$

where ϕ_y is the polar angle of $\hat{\mathbf{y}}$, one sees that the sound field coefficients are given by

$$\beta_n^{(d)}(\omega) = (-i)^n e^{-in\phi_y}.$$

Loudspeaker filter weights are chosen using the least squares approach of Sec. IV D.

Figure 7 illustrates the reproduction of a plane wave approaching from an angle of $\phi_y = \pi/6$. We provide a free field design [Fig. 7(a)], the same free field design in the reverberant room described above [Fig. 7(b)], and the rever-

berant field design [Fig. 7(c)]. With a 307% reproduction error, the reverberant performance of the free-field design is poor. In contrast, the reverberant design performs as well under such conditions as the free-field design does in a free field. Since the -4.4 dB direct-to-reverberant ratio here is common in room environments, we see the importance of reverberant field design techniques.

B. Reproduction of a phantom monopole source

The pressure field of a 2-D monopole source of unit strength is now reproduced. For a monopole source positioned at \mathbf{y} , the sound pressure is

$$P_d(x; \omega) = \mathcal{H}_0^{(2)}(k\|\mathbf{x}-\mathbf{y}\|).$$

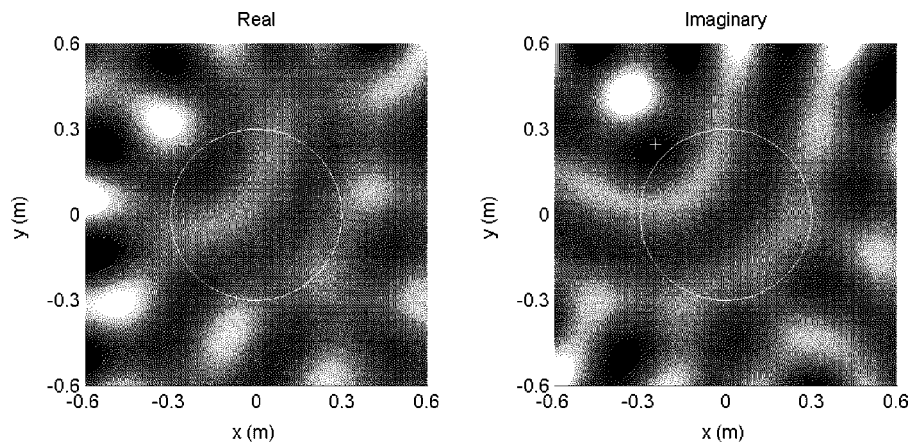


FIG. 9. The reproduction of a plane wave in a 0.6 m radius circle of the reverberant room with 13 speakers and 35 pressure samples when designed over (a) the whole region and (b) a region of 0.3 m radius. Reproduction errors are 26.0% and 84.1%, respectively.

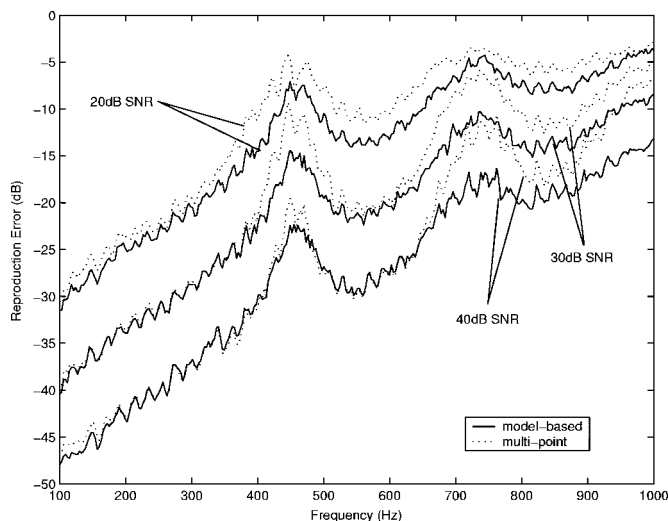


FIG. 10. Wide band reproduction of a plane wave in a 0.3 m radius circle with 13 speakers and using 40 pressure samples, using the model-based method (solid lines) and multiple-point method (broken lines). Reproduction error curves have been averaged over 40 trial runs.

Through the addition property of the Hankel function,²⁰

$$\mathcal{H}_0^{(2)}(k\|\mathbf{x}-\mathbf{y}\|) = \sum_{n=-\infty}^{\infty} \mathcal{H}_n^{(2)}(ky) e^{-in\phi_y} J_n(kx) e^{in\phi_x},$$

where $y \triangleq \|\mathbf{y}\|$, one sees that the sound field coefficients are given by

$$\beta_n^{(d)}(\omega) = \mathcal{H}_n^{(2)}(ky) e^{-in\phi_y}.$$

Using the same design technique as for the plane wave reproduction, we simulate the reproduction of a monopole source just outside the region of interest, at $(y, \phi_y) = (0.35 \text{ m}, 3\pi/4)$. Figure 8 shows a good reproduction of this monopole source.

C. Reproduction with an inadequate numbers of speakers

We now illustrate the result of designs with insufficient numbers of speakers. Again a plane wave is reproduced with 13 speakers, but over a region of interest of radius 0.6 m. At this radius, 25 basis functions are active. For comparison, we show the design for a radius of 0.3 m, where only the 13 lowest-order basis functions are reproduced. Because of the larger radius, we require more pressure samples (35 samples) for these designs.

The resulting sound fields are shown in Fig. 9. While the 0.3 m design reproduces accurately over where the 13 lowest-order basis functions are active [Fig. 9(b)], the 0.6 m design reproduces the sound field with better accuracy over the whole region of interest [Fig. 9(a)].

D. Wide band reproduction with measurement noise

Wide band sound field reproduction of a plane wave is performed with noisy pressure samples in the frequency range 100 Hz to 1 kHz, $R_1=0.3 \text{ m}$, and $R_2=0.27 \text{ m}$. The reproduction error is plotted in Fig. 10 for several noise

SNRs averaged over 40 trial runs. This figure shows that at least 30 dB SNR is required for an accurate reproduction over the whole frequency range.

For comparison the multiple-point method has been co-plotted. As can be seen, the model-based method, by performing the least squares design over the whole region of interest, consistently outperforms the multiple-point method, typically up to 5 dB.

The general trend in this curve is that error increases with frequency. This trend is due to the linear increase in demand for loudspeakers and sensors with frequency. Our design uses the same number of loudspeakers and pressure samples for all frequencies. If we desire to flatten the curve, we could use less pressure samples and loudspeakers at lower frequencies where less basis functions are active.

Also observe the peaks in Fig. 10. These peaks occur in the vicinity of the zeros of the Bessel functions $J_0(kR)$ and $J_1(kR)$. Zeros of these Bessel functions at 460 and 730 Hz, respectively. These peaks are hence a direct result of the error scaling mentioned in Sec. IV C. To flatten such peaks, more pressure sampling should be performed about these frequencies, or the sensor pairs further separated (i.e., $\delta R = R_1 - R_2$ should be increased).

V. CONCLUSION

We have described a novel method of performing sound field reproduction in reverberant enclosures. The key to this method is an efficient parametrization of the acoustical transfer functions. Using this parametrization, we have outlined a practical technique to precisely measure the acoustical transfer functions from a loudspeaker to each point in the region of sound reproduction. This approach allows full sound reproduction without prior knowledge of the loudspeaker positions nor the transmission characteristics of each loudspeaker. Through simulation, the reverberant field method is shown to perform as well in a reverberant room as free-field techniques do in a free field, and up to 5 dB better than multipoint least squares designs. The practical implementation of this soundfield reproduction scheme and its subjective performance remain as open questions and shall be addressed in future research.

APPENDIX A: PROOF OF EQ. (20)

Substituting Eq. (19) into Eq. (18) yields

$$\hat{\beta}_n(\omega) = \frac{1}{J_n(kR)} \frac{1}{M} \sum_{m=0}^{M-1} P(R, \phi_m; \omega) e^{-i(2\pi mn/M)}, \quad (\text{A1})$$

where $\phi_m = 2\pi m/M$. Evaluating the basis function expansion of the sound field Eq. (16) at the M points (R, ϕ_m) :

$$P\left(R, \frac{2\pi m}{M}; \omega\right) = \sum_{q=-\infty}^{\infty} \beta_q(\omega) J_q(kR) e^{i(2\pi mq/M)}. \quad (\text{A2})$$

Substituting Eq. (A2) into Eq. (A1) and interchanging summations:

$$\hat{\beta}_n(\omega) = \frac{1}{J_n(kR)} \frac{1}{M} \sum_{q=-\infty}^{\infty} \beta_q(\omega) J_q(kR) \sum_{m=0}^{M-1} e^{i[2\pi m(q-n)]/M}.$$

Now the summation of the complex exponential is given by

$$\sum_{m=0}^{M-1} e^{i[2\pi m(q-n)]/M} = \begin{cases} M, & \text{if } q-n|M, \\ 0, & \text{otherwise.} \end{cases}$$

Hence

$$\hat{\beta}_n(\omega) = \frac{1}{J_n(kR)} \sum_{s=-\infty}^{\infty} \beta_{n+sM}(\omega) J_{n+sM}(kR).$$

Rearranging,

$$\hat{\beta}_n(\omega) - \beta_n(\omega) = \frac{1}{J_n(kR)} \sum_{s=-\infty, \neq 0}^{\infty} \beta_{n+sM}(\omega) \times J_{n+sM}(kR).$$

QED.

APPENDIX B: BOUND ON THE TERMWISE SCALING FACTOR κ_n

For the following discussion, we view $J_n(kR)$ as a function of its order n . For $n \geq N = \lceil kR \rceil$, $J_n(kR)$ is observed to be a monotone decreasing function in n , decaying exponentially toward zero. (This property can be observed in Fig. 5 for n up to 20.) Similarly for $n \geq N$, $|J_{-n}(kR)|$ is also monotone decreasing in n .

From Eq. (21), the largest termwise scaling factor for coefficient $\hat{\beta}(\omega)$ is

$$\kappa_n = \frac{1}{J_n(kR)} \max_{q=-\infty, \dots, \infty, \neq 0} |J_{n+qM}(kR)|.$$

Now since $M > 2N$, for $n = -N, -N+1, \dots, N$ and $q = \pm 1, \pm 2, \dots$, we have $|n+qM| > N$. Each of these Bessel functions $J_{n+qM}(kR)$ is hence sampled over the above-mentioned monotone decreasing interval. Consequently, κ_n is maximized when $|n+qM|$ is minimized:

$$\kappa_n = \begin{cases} J_{-(M-n)}(kR)/J_n(kR), & n \geq 0, \\ J_{M+n}(kR)/J_n(kR), & n < 0. \end{cases}$$

Calculating now the maximum κ_n over the active basis functions of positive index, the termwise scaling factor is bounded by

$$\begin{aligned} \max_{n=0, \dots, N} \kappa_n &= \max_{n=0, \dots, N} \left| \frac{J_{-(M-n)}(kR)}{J_n(kR)} \right| \\ &\leq \max_{n=0, \dots, N} 1/|J_n(kR)| \times \max_{n=0, \dots, N} |J_{-(M-n)}(kR)|. \end{aligned}$$

Again, due to the monotone decreasing property of the Bessel function:

$$\max_{n=0, \dots, N} \kappa_n = \max_{n=0, \dots, N} 1/|J_n(kR)| \times |J_{-(M-N)}(kR)|.$$

Similarly, due to the Bessel function property $J_{-n}(x)$

$= (-1)^n J_n(x)$, the bound is the same for $\max_{n=-N, \dots, -1} \kappa_n$.

Then applying the Bessel function bound from Ref. 18,

$$\max_{n=-N, \dots, N} \kappa_n = \max_{n=0, \dots, N} 1/|J_n(kR)|$$

$$\times \frac{1}{\sqrt{2\pi(M-N)}} \left[\frac{ekR}{2(M-N)} \right]^{M-N}.$$

QED.

¹M. A. Gerzon, "Ambisonics in multichannel broadcasting and video," *J. Audio Eng. Soc.* **33**, 859–871 (1985).

²R. Nicol and M. Emerit, "3d-sound reproduction over an extensive listening area: a hybrid method derived from holophony and ambisonic," in *AES 16th International Conference on Spatial Sound Reproduction*, Helsinki, 1999, Vol. II, pp. 436–453.

³M. A. Poletti, "A unified theory of horizontal holographic sound systems," *J. Audio Eng. Soc.* **48**, 1155–1162 (2000).

⁴A. J. Berkout, D. D. Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.* **93**, 2764–2778 (1993).

⁵S. Takane, Y. Suzuki, and T. Sone, "A new method for global sound field reproduction based on Kirchhoff's integral equation," *Acustica* **85**, 250–257 (1999).

⁶S. Ise, "A principle of sound field control based on the Kirchhoff–Helmholtz integral equation and the theory of inverse systems," *Acustica* **85**, 78–87 (1999).

⁷O. Kirkeby and P. A. Nelson, "Reproduction of plane wave sound fields," *J. Acoust. Soc. Am.* **94**, 2992–3000 (1993).

⁸O. Kirkeby, P. A. Nelson, F. Orduna-Bustamante, and H. Hamada, "Local sound field reproduction using digital signal processing," *J. Acoust. Soc. Am.* **100**, 1584–1593 (1996).

⁹D. B. Ward and T. A. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Trans. Speech Audio Process.* **9**, 697–707 (2001).

¹⁰J. Mourjopoulos, "On the variation and invertibility of room impulse response functions," *J. Sound Vib.* **102**, 217–228 (1985).

¹¹P. A. Nelson and F. Orduna-Bustamante, "Inverse filter design and equalization zones in multichannel sound reproduction," *IEEE Trans. Speech Audio Process.* **3**, 185–192 (1995).

¹²O. Kirkeby and P. A. Nelson, "Digital filter design for inversion problems in sound reproduction," *J. Audio Eng. Soc.* **47**, 583–595 (1999).

¹³J. Mourjopoulos, "Digital equalization of room acoustics," *J. Audio Eng. Soc.* **42**, 884–900 (1994).

¹⁴S. Bharitkar and C. Kyriakakis, "A cluster centroid method for room response equalization at multiple locations," in *Proceedings of the IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, Mohonk, 2001.

¹⁵F. Asano and D. C. Swason, "Sound equalization in enclosures using modal reconstruction," *J. Acoust. Soc. Am.* **98**, 2062–2069 (1995).

¹⁶A. O. Santillan, "Spatially extended sound equalization in rectangular rooms," *J. Acoust. Soc. Am.* **110**, 1989–1997 (2001).

¹⁷E. G. Williams, *Fourier Acoustics* (Academic, London, 1999).

¹⁸H. M. Jones, R. A. Kennedy, and T. D. Abhayapala, "On dimensionality of multipath fields: spatial extent and richness," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Orlando, 2002, Vol. III, pp. 2837–2840.

¹⁹J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.* **65**, 943–950 (1979).

²⁰D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory* (Springer-Verlag, Berlin, 1992).

Microphone array signal processing with application in three-dimensional spatial hearing

Mingsian R. Bai^{a)} and Chenpang Lin

Department of Mechanical Engineering, National Chiao-Tung University, 1001 Ta-Hsueh Road, Hsin-Chu 300, Taiwan, Republic of China

(Received 3 August 2004; revised 4 December 2004; accepted 6 December 2004)

Microphone arrays are known to enhance the directionality and signal-to-noise ratio (SNR) over single-channel sensors. This is considered beneficial in many applications such as video-conferencing systems and hearing aids. However, this advantage comes at the price of the sensation of spatial hearing. The spatial cues due to diffractions of the head and torso are lost if the array is not fitted in the ears. In this paper we present a system that incorporates binaural hearing synthesis into array signal processing, in an attempt to recover the three-dimensional sound image that a human listener would naturally perceive. In the system, the superdirective beamformer is exploited to estimate the direction of arrival (DOA) of the incoming sound. The spatial sound image is restored by steering the beam to the direction found in the DOA session and filtering the array output with the corresponding Head Related Transfer Functions (HRTF). The algorithms have been implemented in real-time fashion using a digital signal processor. Objective and subjective experiments were performed to validate the proposed system. The experimental results showed that the accurate localization of the sound source is achievable using the array system. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1853242]

PACS numbers: 43.60.Fg [EJS]

Pages: 2112–2121

I. INTRODUCTION

Microphone arrays have received much research interest as a means of acoustic pickup utilized in various applications such as video-conferencing systems¹ and hearing aids.^{2–5} One of the reasons for using an array is to improve signal-to-noise ratio (SNR) that has long been a plaguing problem of, for example, conventional single channel hearing aids. This problem is further aggravated in the environments where reverberations and interferences are present. In comparison with single-channel sensors, microphone arrays provide advantages that the SNR as well as directionality of the sensor can be enhanced using such a system. In particular, an array behaves like a spatial filter, enabling the listener to focus on the signal source such as speech and at the same time reject ambient noise and interference.⁶ This is an attractive feature for hearing-impaired people who may desire low-noise hearing aids. A great number of array signal processing methods can be found in the literature⁷ to design arrays subject to individual requirements in an application. Beamforming and estimation of direction of arrival (DOA) are known to be two major functions of arrays. These two functions serve to track the intended source in a particular direction with high signal quality.

However, the above-mentioned advantages come at the price of sensation of spatial hearing. The spatial cues due to diffractions of the head and torso are lost if the sensors are not fitted in ears. To address the problem, in this paper we present a system that incorporates binaural hearing synthesis into array signal processing, in an attempt to recover the three-dimensional (3-D) sound image that a human listener

would naturally perceive. In order to form a very sharp beam, the superdirective array^{8–13} is employed in the array design. The thus designed beam is then electronically steered in every direction to estimate the DOA. Once the DOA is found, the sound beam is fixed at that direction. Finally, the binaural signals of the spatial sound image are produced by filtering the array output with the corresponding Head Related Transfer Functions (HRTF).¹⁴ The algorithms are implemented in real-time fashion using a digital signal processor, TMS320C32. As such, a low noise sound with high spatial quality is reproduced with the aid of the thus integrated array–HRTF system.

Simulations and experiments are carried out to evaluate the proposed system. The SNR gain and the effects of aperture size on the directivity of the microphone arrays are examined. Objective and subjective tests were performed to see how effective the listeners localize the source of sound using the proposed array system.

II. MICROPHONE ARRAY SIGNAL PROCESSING

It is well known that microphone arrays are capable of enhancing the directionality and signal-to-noise ratio (SNR) over single-channel sensors. This is beneficial in many applications such as hearing aids. In this paper, array signal processing techniques are utilized for the DOA estimation and beamforming. DOA estimation refers to localizing the source direction, while beamforming refers to forming a beam pattern with a specified shape and orientation. In this section, two algorithms will be presented to carry out these tasks of array signal processing, followed by an investigation on the directivity pattern.

^{a)}Corresponding author. Electronic mail: msbai@mail.nctu.edu.tw

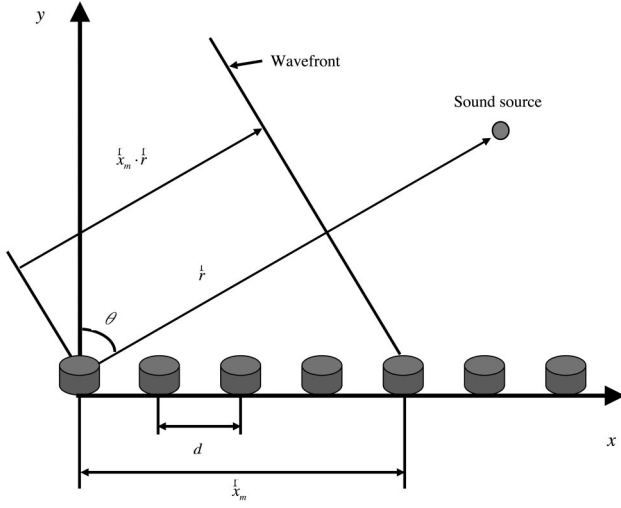


FIG. 1. The configuration of a uniform linear array (ULA). A sound source is located in the farfield. \vec{x}_m is the position vector of the m th sensor and \vec{r} is the position vector of the farfield source.

A. Array signal processing algorithms

Figure 1 shows a uniform linear array (ULA), in which sensors are allocated along a straight line with equal spacing d . It is assumed that the sound source is at the farfield and the signals received at the sensors are narrowband with center frequency ω . The spacing between adjacent microphones is less than one-half wavelength to avoid grating lobes.

1. The delay-sum array

A brief review of the delay-sum method is given as follows. For an M -sensor array, the signals received at the sensors, $x_1(t), \dots, x_M(t)$, form the *data vector*

$$\mathbf{x}(t) = \mathbf{a}(\vec{r})s(t) + \mathbf{n}(t), \quad (1)$$

where $\mathbf{a}(\vec{r}) = [e^{j\omega[(\vec{r}_1 \cdot \vec{r})/c]} \dots e^{j\omega[(\vec{r}_M \cdot \vec{r})/c]}]^T$ is termed the *array manifold vector*, \vec{r} is the unit vector pointing to the source, \vec{r}_m , $m = 1, \dots, M$ are the position vectors of the sensors, c is the speed of sound, $s(t)$ represents the baseband signal of the source and $\mathbf{n}(t) = [n_1(t) \dots n_M(t)]^T$ is the noise vector. The beamformer output is the weighted sum of the delayed input signals, given by

$$y(t) = \mathbf{w}^H \mathbf{x}(t) = \mathbf{w}^H \mathbf{a}(\vec{r})s(t) + \mathbf{w}^H \mathbf{n}(t), \quad (2)$$

where $\mathbf{w} = [w_1 \dots w_M]^T$ is the array weight vector and the operator H denotes the complex conjugate transposition. The delay-sum algorithm is to “time align” the received signals by choosing the following weight vector:

$$w_m = \frac{1}{M} e^{-j\omega[(m-1)d \sin \theta/c]} = e^{-j\omega\tau_m}, \quad m = 1, \dots, M, \quad (3)$$

where θ is the look angle (measured from the normal) of the array and is dependent of the source vector \vec{r} and

$$\tau_m = \frac{(m-1)d \sin \theta}{c} \quad (4)$$

is the delay that the m th channel needs to compensate. It can be shown that the delay-sum algorithm attains the maximum signal-to-noise ratio gain (SNRG).⁷ The delay τ_m usually is

TABLE I. FIR filter coefficients for the first-order ($N=1$) and the second-order ($N=2$) Lagrange interpolation.

	w_{m0}	w_{m1}	w_{m2}
$N=1$	$1 - e_m$	e_m	
$N=2$	$(e_m - 1)(e_m - 2)/2$	$-e_m(e_m - 2)$	$e_m(e_m - 1)/2$

not an integer. The simplest approach to deal with these fractional delays is the Lagrange interpolation method.¹⁵ We first divide τ_m by the sampling period T to acquire the fractional delay that can be written into the integer and fractional components, D_m and e_m , respectively,

$$\frac{\tau_m}{T} = D_m + e_m. \quad (5)$$

The Finite Impulse Response (FIR) filter coefficients required to realize the fractional delay are given by

$$w_{mk} = \prod_{\substack{l=0 \\ l \neq k}}^N \frac{e_m - l}{k - l}, \quad k = 0, 1, 2, \dots, N. \quad (6)$$

The coefficients for the Lagrange filters of order $N=1, 2$ are given in Table I. The expectation value of $|y(t)|^2$ plotted versus the look angle θ is called the spatial power spectrum:

$$S(\theta) = E\{|y(t)|^2\} = \mathbf{w}^H \mathbf{R}_{xx} \mathbf{w}, \quad (7)$$

where $\mathbf{R}_{xx} \triangleq E\{\mathbf{x}(t)\mathbf{x}^H(t)\}$ is the *data correlation matrix*. The peak of the spatial power spectrum corresponds to the direction of the sound source.

2. The superdirective array

Another array algorithm employed in this work is the superdirective array.¹³ The principle of this algorithms follows from maximizing the *array gain* that is the measure of improvement of the signal-to-noise ratio between one sensor and the array output,

$$G = \frac{\text{SNR}_{\text{Array}}}{\text{SNR}_{\text{Sensor}}}. \quad (8)$$

The larger the array gain, the higher the ability of the array as a spatial filter to suppress the noise. The above array gain can be shown to be equivalent to¹³

$$G = \frac{|\mathbf{w}^H \mathbf{d}|^2}{\mathbf{w}^H \mathbf{\Gamma}_{nn} \mathbf{w}}, \quad (9)$$

where $\mathbf{\Gamma}_{nn}$ is the *coherence matrix* of the noise and \mathbf{d} is the steering vector of the main axis of the array. It is generally assumed that a diffuse white noise field is spherically isotropic, in which the ab th entry of the matrix takes the form¹³

$$\Gamma_{n_a n_b} = \frac{\sin[k(a-b)d]}{k(a-b)d}. \quad (10)$$

On the other hand, if only the self-noise of the sensors is present, i.e., $\mathbf{\Gamma}_{nn} = \mathbf{I}$, the array gain G reduces to the white noise gain (WNG):

$$\text{WNG} = \frac{|\mathbf{w}^H \mathbf{d}|^2}{\mathbf{w}^H \mathbf{w}}. \quad (11)$$

In general, the larger the array weights, the smaller the WNG and hence the more sensitive to noise is the array. Another important quantity to evaluate arrays, the directivity index (DI) is the logarithmic equivalent of the above-mentioned array gain,

$$DI = 10 \log_{10} \left(\frac{|\mathbf{w}^H \mathbf{d}|^2}{\mathbf{w}^H \mathbf{\Gamma}_{nn} \mathbf{w}} \right), \quad (12)$$

which amounts to the ratio of the main-axis gain over the angle-weighted gain. The larger the DI, the more directional is the array pattern.

The idea of the superdirective design is realized by maximizing the DI in Eq. (12), or equivalently, the array gain in Eq. (8). Equation (8) is in fact a Rayleigh's quotient. The maximization of which can be achieved by solving the following optimization problem:

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{\Gamma}_{nn} \mathbf{w}, \quad \text{subject to } \mathbf{w}^H \mathbf{d} = 1. \quad (13)$$

In other words, we look for an optimal weight \mathbf{w} such that the array output power is minimized with the gain at the look direction constrained to unity. Thus, the array aims to receive an undistorted signal response at the main axis and reject unwanted interferences at the other directions. Following the method of the Lagrange multiplier, the solution of Eq. (13) is given as^{7,16}

$$\mathbf{w} = \frac{\mathbf{\Gamma}_{nn}^{-1} \mathbf{d}}{\mathbf{d}^H \mathbf{\Gamma}_{nn}^{-1} \mathbf{d}}. \quad (14)$$

In the low-frequency range, the matrix $\mathbf{\Gamma}_{nn}$ is nearly singular. The direct inverse of $\mathbf{\Gamma}_{nn}$ would prove problematic and results in exceedingly large array weights. To address the problem, a simple regularization procedure is generally utilized by incorporating a positive constant to the main diagonal of the coherence matrix:¹³

$$\mathbf{w}|_{\text{regularized}} = \frac{(\mathbf{\Gamma}_{nn} + \epsilon \mathbf{I})^{-1} \mathbf{d}}{\mathbf{d}^H (\mathbf{\Gamma}_{nn} + \epsilon \mathbf{I})^{-1} \mathbf{d}}. \quad (15)$$

The parameter ϵ can vary anywhere from zero to infinity, which corresponds to the unconstrained superdirective array or the delay-sum array, respectively. For instance, a reasonable value of ϵ to compromise the array directivity and the weight size is 0.01.

The delay-sum method and the superdirective method of a broadside array ($\theta=0^\circ$) are compared in terms of DI, WNG, and optimal weights as follows. Assume that there are 4 sensors equally spaced with 4 cm. Figure 2(a) illustrates the DI of the delay sum and the superdirective arrays with $\epsilon=0.01$. Clearly visible is the improvement of DI below 3 kHz achieved using the superdirective design over the delay-sum method. Figure 2(b) compares the WNG of the delay sum and the superdirective methods with $\epsilon=0.01$ for the broadside array. The WNG of the delay-sum array is larger than the superdirective method at the expense of low-frequency directionality. Figure 2(c) compares the 2-norm of the weight vectors of the delay sum and the superdirective methods with $\epsilon=0.01$ for the broadside array. It can be seen from the result that the optimum weights of the superdirec-

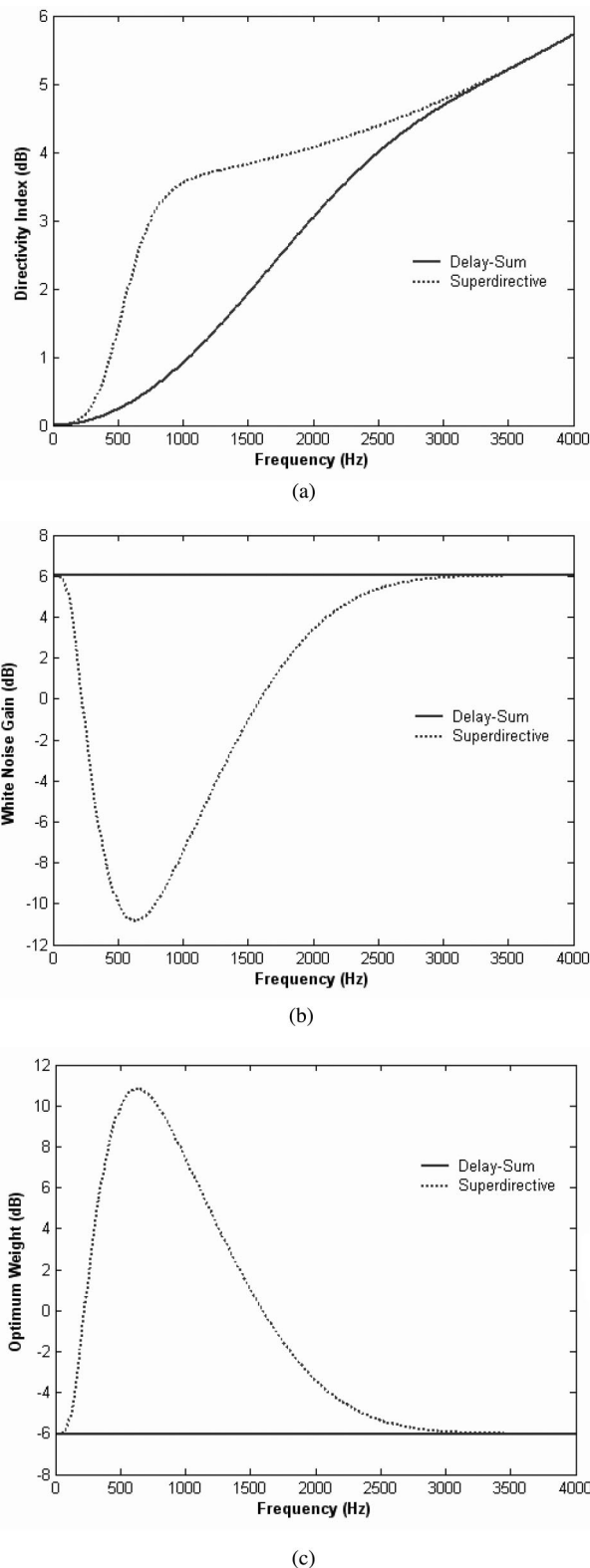


FIG. 2. A comparison of the delay-sum and the superdirective arrays with $\epsilon=0.01$. (a) Directivity index. (b) White noise gain. (c) Optimum weight norm.

tive array are larger than the delay-sum array at low frequencies in order to attain high directionality. The weights of the superdirective array should be maintained under a certain level in order not to create problems of filter implementation. To see more details, Fig. 3 shows the contour plots of direc-

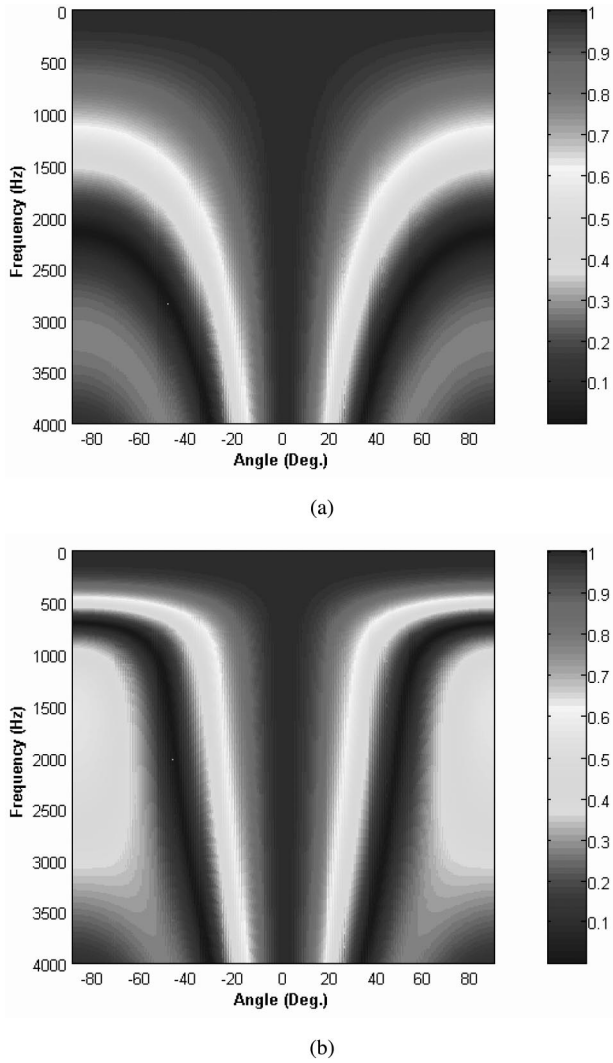


FIG. 3. The contour plots (in a linear scale) of the directional responses versus angle (x axis) and frequency (y axis) of a four-element broadside array. The interelement spacing is 4 cm. (a) Delay-sum array (b) Superdirective array.

tivity versus the angle and frequency of the delay sum and superdirective arrays. The directivity is significantly improved using the superdirective method in the band 500–1500 Hz.

The foregoing analysis based on the narrowband formulation can now be extended to the broadband scenario. The array weights obtained using the superdirective method is associated with the frequency ω . Repeating this process for the frequencies equally spaced in the band of interest gives the frequency response samples of each array filter. Now that the frequency response samples are obtained, the inverse discrete Fourier transform is applied to acquire the impulse response, or the filter coefficients, of the superdirective array. More precisely, if the P_w frequency samples of the array weight are obtained for the m th array filter, the discrete frequency response of the filter is simply

$$H_m(l) = w_m^*(l), \quad l = 1, \dots, P_w. \quad (16)$$

To assure real impulse responses, the frequency response samples $H_m(l)$ at $l = -1, \dots, -(P_w - 1)$ must be mirrored to

$l = P_w + 1, \dots, 2P_w - 1$ with proper symmetry. The FIR filter coefficients can be obtained by applying the inverse discrete Fourier transform (IDFT) to the frequency response samples for each channel:

$$h_m(k) = \frac{1}{2P_w} \sum_{l=0}^{2P_w-1} H_m(l) W_{2P_w}^{-lk},$$

$$k = 1, \dots, (2P_w - 1); \quad m = 1, \dots, M, \quad (17)$$

where $W_{2P_w} = \exp[-j(\pi/P_w)]$. The thus obtained $h_m(k)$ is often noncausal. A circular shift with one-half the IDFT length can be performed to allow for a causal filter. Assume that there are 4 sensors equally spaced with 4 cm. The impulse responses of the superdirective array implemented using 64-tap FIR filters are shown in Fig. 4(a). The frequency responses of the superdirective filters are shown in Figs. 4(b) and 4(c). The impulse responses are symmetric and exhibit sign flipping between the microphones 1 and 2, and also the microphones 3 and 4. The frequency responses show an even more pronounced high gain and phase switching at the low-frequency range. These interesting phenomena suggest that the differential actions are necessary to produce the directivity improvement for superdirective microphones. Nevertheless, the superdirective microphones do not suffer from the poor SNR problem due to the ω^n dependence, as do pure differential microphones.¹⁷

Although there are many techniques available for DOA determination, we choose the simplest but most robust approach, the Fourier beamforming, due to the concern of computational loading of our processor. In the DOA session, angle spectra are calculated by steering the main beam from -90° to 90° using fractional delays.¹⁵ The maxima of angle spectra correspond to the source direction in the farfield. Figure 5 shows the angle spectra estimated at 500, 1000, 2000, and 4000 Hz by using the delay-sum and the superdirective methods, respectively. The result is plotted in a linear scale to facilitate the determination of DOA from the maxima of the angle spectra. The sound source is oriented at 10° , as indicated by the maxima of the angle spectra. The superdirective method produces a sharper DOA estimation than the delay-sum method at 500, 1000, and 2000 Hz. At 4000 Hz and higher, the superdirective method has no particular advantage. However, the side lobes of the superdirective array are larger than the delay-sum array, which is the price to pay for better resolution in DOA estimation. There is obviously a tradeoff between the beamwidth and the side lobes in the array design.

B. Performance analysis of array

Directivity is an important feature of microphone arrays, which is highly dependent on the aperture size of the arrays. In what follows, simulations and measurements were carried out to examine the directivity pattern of a 4-element linear microphone array. The signals from the microphones are summed directly without any filtering. The interelement spacing is 4 cm and the total length (aperture size) of microphone array is 12 cm. The sound source is positioned at 0°

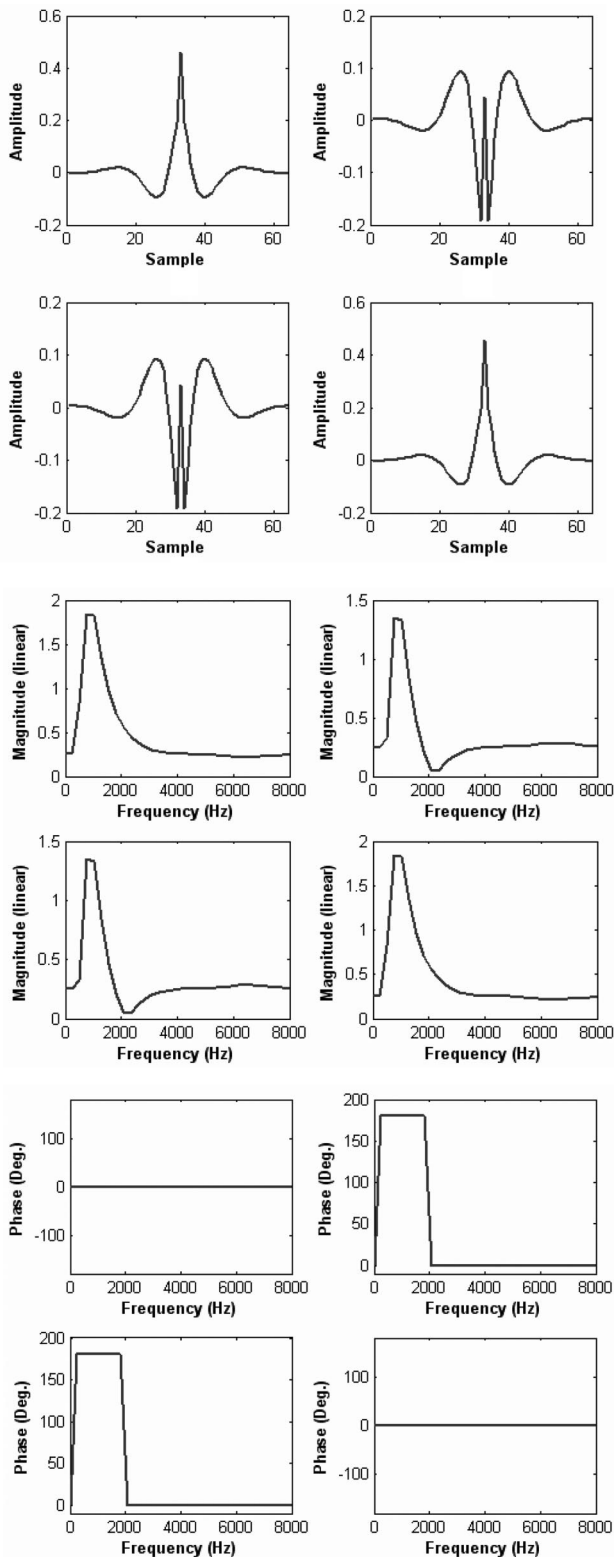


FIG. 4. The characteristics of the superdirective filter. The filter is realized as a 64-tapped FIR filter. The results are shown as 2×2 matrices, where the (1,1), (1,2), (2,1) and (2,2) correspond to the response of the 1st, 2nd, 3rd and 4th sensor, respectively. (a) The impulse response. (b) The frequency response magnitude. (c) The frequency response phase.

For reference, the simulation and measurement results of the directivity patterns of one microphone are shown in Fig. 6. The directivity pattern of one microphone at all frequencies in the simulation is omnidirectional. The directivity pat-

tern at 500 and 1000 Hz of one microphone in the measurement are also omnidirectional, while the measurements at 2000 and 4000 Hz display slight directivity due to the effect of the baffle in which the microphones are embedded.

Figure 7 shows the simulation and measurement results of the directivity patterns of the four-element array. The directivity patterns at 500 and 1000 Hz in the simulation are nearly omnidirectional, while the directivity pattern is quite directional at 2000 and 4000 Hz. The measured data are in close agreement with the simulation. These results indicate that the direct-sum array would display directivity only above 2000 Hz. More sophisticated algorithms must be used to produce significant directivity at a low-frequency range.

Apart from directivity pattern, the SNRs of one microphone and the four-element array are also measured and listed in Table II. The SNR gain improved by a 12 cm aperture microphone array is 11.6 dB, which is close to the theoretical value of 12 dB. This suggests that the SNR can be enhanced over a single sensor by means of array structures.

III. SPATIAL SOUND RESTORATION USING HRTF

As mentioned earlier, microphone arrays have the advantage of improved SNR, directivity, and hence spatial selectivity. These features help the rejection of undesirable effects of room reverberation and acoustic feedback. However, these benefits of arrays can be countered by the loss of three-dimensional (3-D) hearing when applied to hearing aids. To address the problem, a post-filtering technique based on the Head Related Transfer Functions (HRTFs) is presented in the paper in an attempt to restore the 3-D spatial hearing mechanism resulting from the head and torso diffractions of humans. Therefore, using this microphone-HRTF system, hearing-impaired people can hear more spatial-sounding speech and music signals with a better source localization.

A. The method and system architecture

An HRTF is a measurement of the transformation for a specific source direction relative to the head, and describes the filtering process associated with the diffraction of sound by the torso, the head, and the pinnae. HRTFs contain important cues, including the interaural time differences (ITD), the interaural level differences (ILD), and spectral characteristics pertaining to spatial hearing and sound localization. ITD refers to the time difference between the left and right ears for a plane wave incident from a certain direction. ILD refers to the level difference due to the head shadowing effect between the levels of the signals received at both ears. The inverse Fourier transform of a HRTF is termed the head-related impulse responses (HRIR). A 3-D sound field can be created by convolving a source signal with the appropriate HRTF to render sound images positioned arbitrarily around a listener. The HRTF and HRIR database employed in the paper is currently available on the web.¹⁴

The entire system consists of a DOA estimation module, a beamsteering module, and a HRTF post-processing module. The underlying idea of the system will be explained as follows. Rewrite the array output in Eq. (2) with a discrete-time and broadband setting,

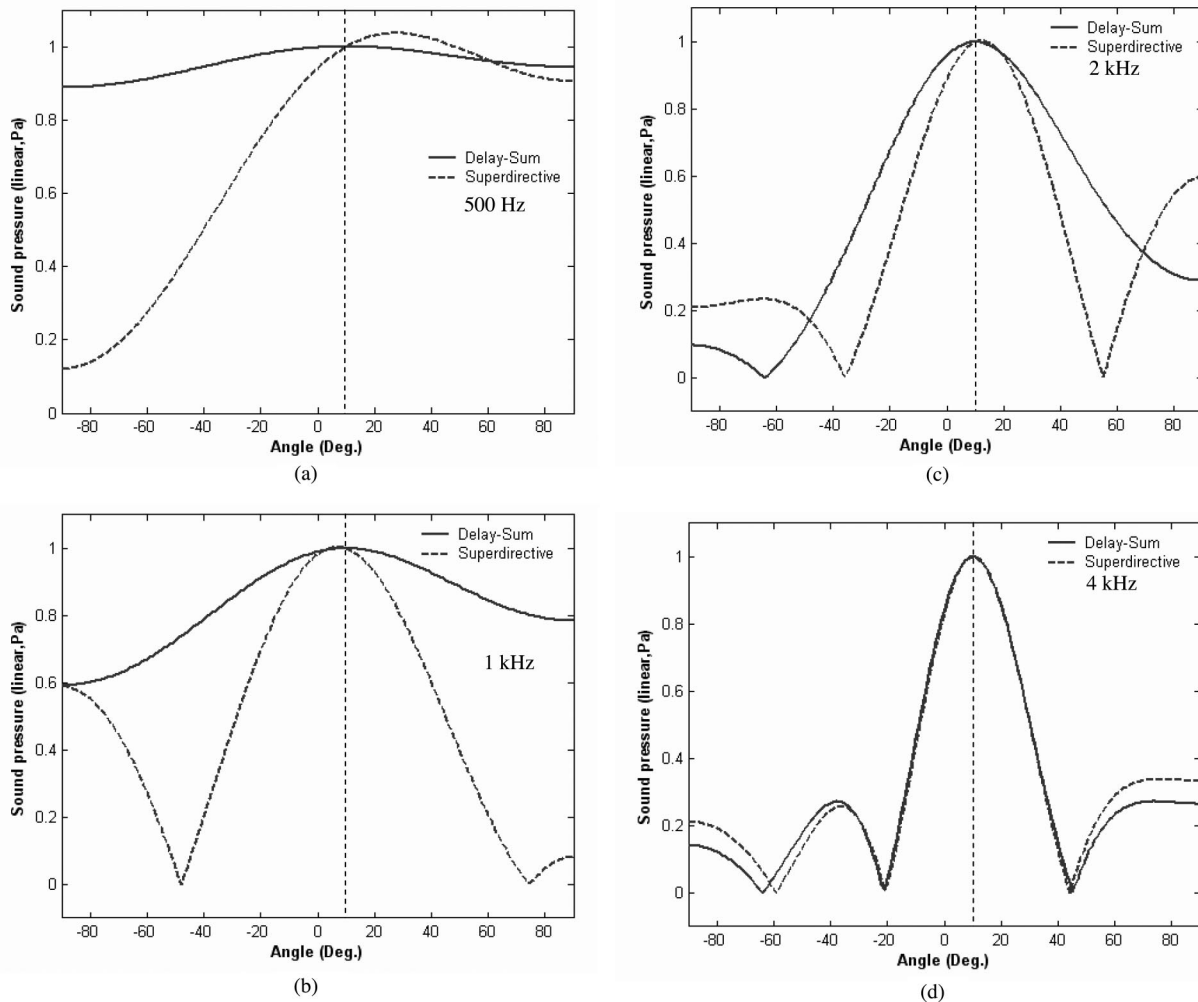


FIG. 5. The simulation results of the DOA estimation (in a linear scale) by the delay-sum and superdirective methods. The sound source is a pure tone and is located at $\theta=10^\circ$, as indicated by a vertical dotted line in the figure. (a) 500 Hz; (b) 1000 Hz; (c) 2000 Hz; (d) 4000 Hz.

$$y(n) = \mathbf{w}^H(n) * \mathbf{x}(n), \quad (18)$$

where $*$ denotes convolution and n is the discrete-time index. In Eq. (17),

$$\mathbf{x}(n) = s(n) * \mathbf{a}(\theta, n) \quad (19)$$

being the data vector received at the microphones, and $\mathbf{a}(\theta, n)$ is the impulse response of the array manifold vector associated with the source direction θ . After a DOA session using the superdirective beamformer, a potential source direction is found. Then, the beam of the array is electronically steered to that direction using the second-order Lagrange interpolation for the steering vector $\mathbf{d} \approx \mathbf{a}$. Note that the beamsteering process should guarantee $\mathbf{w}^H \mathbf{d} \approx 1$. Although this is a fixed beamformer, we choose the simple but practical approach for two reasons. First, apart from processing needs in multichannel array filtering, DOA, HRTF, and real-time display, there is really not much computational power left in the present DSP system to afford sophisticated adaptive algorithms such as the generalized sidelobe canceller (GSC) or the Griffiths beamformer.⁷ Second, adaptive beamformers still have robustness issues in the context of a steering vector error and correlated noise.¹³ In particular, the latter issue frequently occurs in a live room environment such as a car

cabinet, where multipath reflections could lead to an undesirable signal-noise cancellation.

In order to recover the spatial sound image for the array system, as a final step, HRIRs are used to filter the array output to yield the binaural signals for a headphone,

$$\mathbf{y}(n) = \begin{bmatrix} h_L(n) \\ h_R(n) \end{bmatrix} * y(n), \quad (20)$$

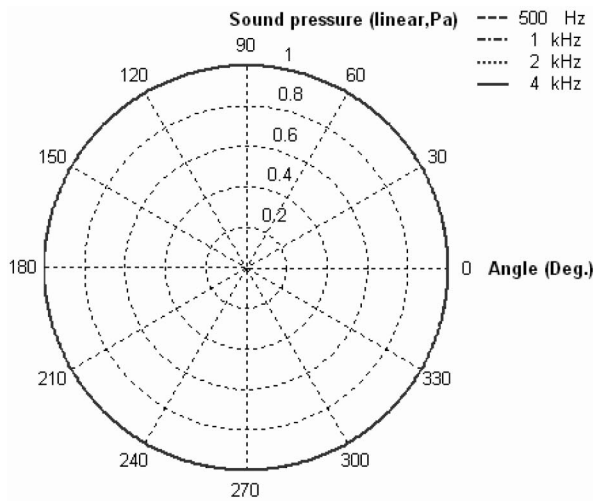
where $h_L(n)$ and $h_R(n)$ are the HRIRs of the left and the right ears. In effect, the overall output that the system will produce is

$$\mathbf{y}(n) = \begin{bmatrix} h_L(n) \\ h_R(n) \end{bmatrix} * s(n) * \mathbf{w}^H(n) * \mathbf{a}(\theta, n). \quad (21)$$

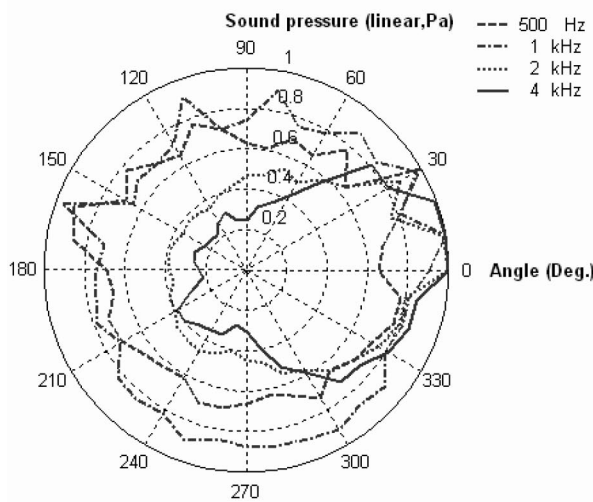
The preceding procedure of post-processing for 3-D hearing is summarized in the flow chart of Fig. 8.

B. Experimental verifications

The above-mentioned array signal processing algorithms for 3-D spatial sound restoration with HRTF was implemented on a digital signal processor (DSP), TMS320C32. Figure 9 shows the system block diagram, including a computer equipped with the DSP, a sound source, a four-element



(a)

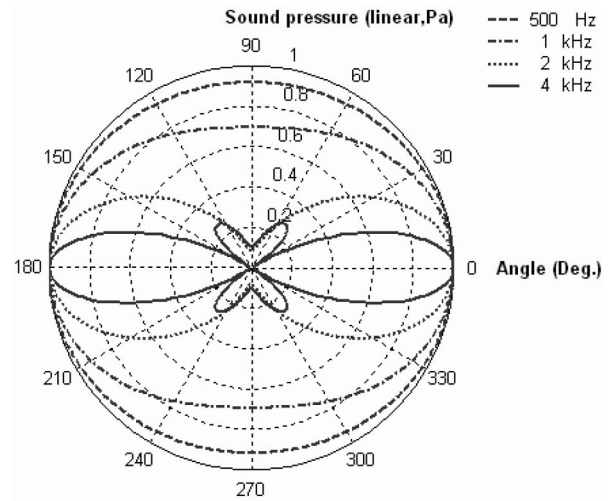


(b)

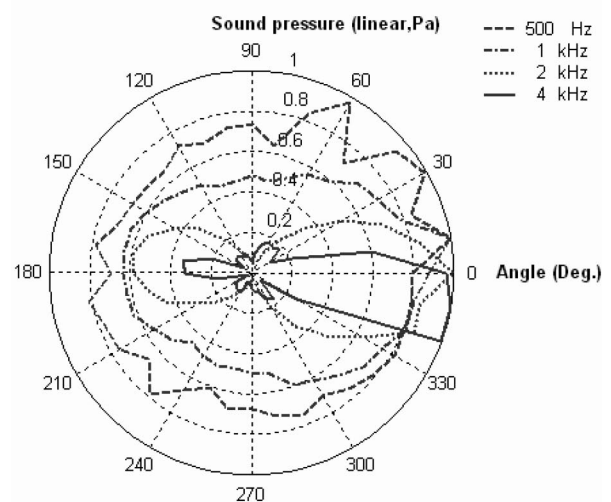
FIG. 6. The directivity pattern of one microphone. (a) Simulation. (b) Measurement.

linear microphone array, a preamplifier, and two second-order antialiasing filters with cutoff frequency 8 kHz and a headphone. Four condenser microphones with 6 mm diameter are fitted in a acrylic plastic plate and the interelement spacing of microphones is 4 cm, which gives a total length of 12 cm. The sound source is placed at $\theta=10^\circ$. The sampling rate was chosen to be 16 kHz.

The delay-sum beamformer and the superdirective beamformer are applied to estimate DOA using the DSP system. Figure 10 shows the experimental results of the DOA estimation of obtained using the delay sum and the superdirective methods at 500, 1000, 2000, and 4000 Hz, respectively. A more accurate DOA estimation at 500, 1000, and 2000 Hz can be obtained by using the superdirective method than the delay-sum method because of the narrower beamwidth of the former. This observation is in agreement with the simulation results. At 4000 Hz, two beamformers yielded a comparable performance. Figure 11 shows the DOA estimation obtained using the delay-sum and the superdirective methods, in which case a sound source located at $\theta=10^\circ$ is a



(a)



(b)

FIG. 7. The directivity pattern of a broadside delay-sum array with four microphones equally spaced by 4 cm. (a) Simulation. (b) Measurement.

random noise bandlimited to 8 kHz. An error analysis of DOA estimation obtained experimentally by the delay-sum beamformer and the superdirective beamformer is illustrated in Table III. The average DOA error of the superdirective beamformer is 0° vs 6° obtained using the delay-sum beamformer. The former method significantly outperforms the latter one, owing to its narrower beamwidth (angle resolution).

After the DOA was estimated, the beamformer in conjunction with beam steering was utilized to enhance the sound signals. Only the superdirective beamformer was applied in the following tests because of its superior angle resolution. The fractional delays should be carefully compensated using Lagrange interpolation. Otherwise, a large

TABLE II. The measured SNRs and SNR gain of the single microphone versus the four-element microphone array with a 12 cm aperture.

Aperture size	Microphone number	SNR (dB)	SNR gain (dB)
12 cm	One	46.1	11.6
	Four	57.7	

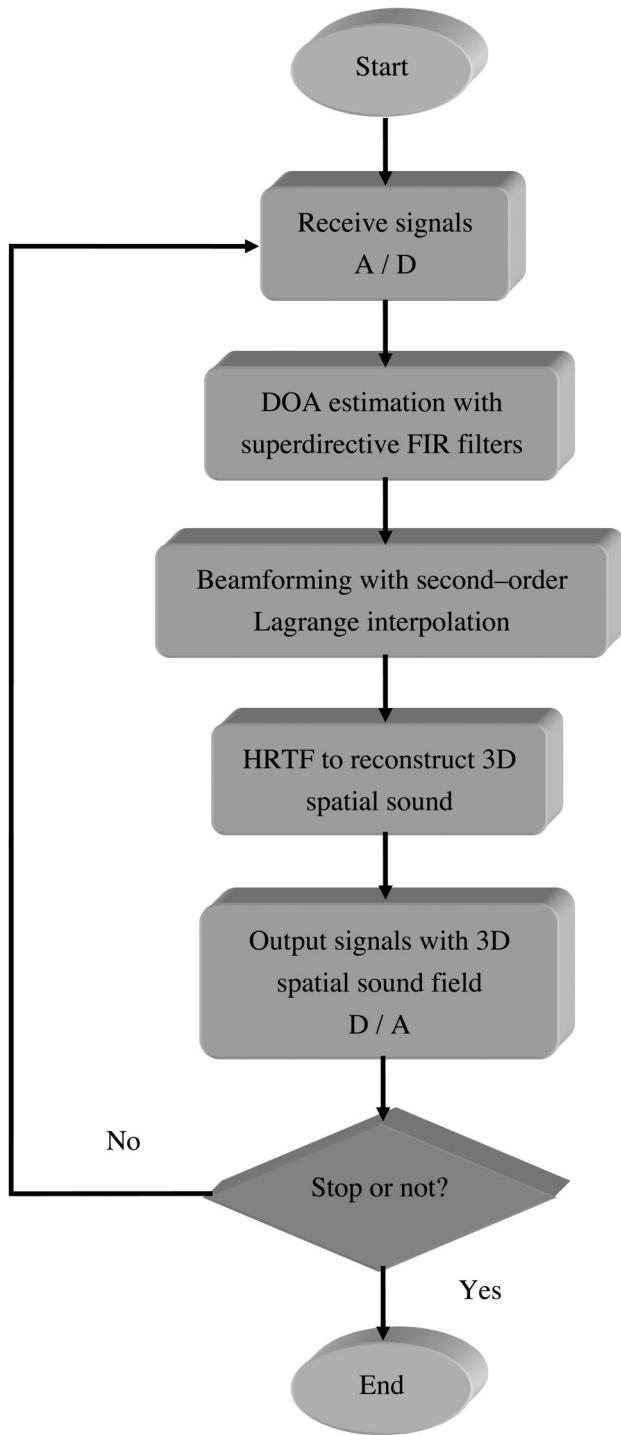


FIG. 8. The flow chart of array signal processing with 3-D spatial sound restoration with the HRTF.

steering error would arise, especially when the sampling rate is low.

Following the beam-steering session, post-processing based on HRTFs was carried out to realize the 3-D spatial sound. When the location of the sound source is changed, HRTFs were also updated according to the new direction found in the DOA session. In order to evaluate the effectiveness of the system in localizing sound sources, a subjective experiment was conducted. The listening test involved ten human subjects. Random noise bandlimited to 8 kHz was

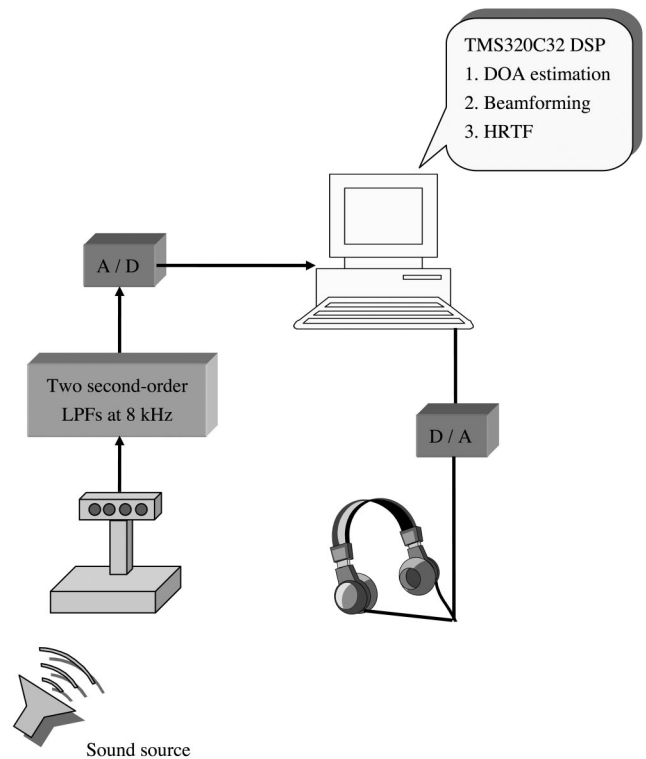


FIG. 9. The block diagram of the array spatial sound system implemented on a DSP, TMS320C32.

used as the source. A headphone was employed as the rendering device. In the experiment, 11 positions, ranging from -50° to 50° with 10° intervals, were preselected to position the source. The distance between the microphone array and the source was 2 m. The experiment was conducted in an anechoic chamber to minimize unwanted reflections. Figure 12 shows the result of the subjective localization experiment. The perceived angles of source are in very good agreement with the presented angles of the source since most data points fall on the diagonal of the plot. Figure 13 shows the statistics of localization errors in a bar chart. It can be seen in the result that the localization errors at $\pm 10^\circ$ and $\pm 50^\circ$ are somewhat larger than the other angles. Overall, the average localization error is 11.7° (nearly one test interval), and the standard deviation is 6.5° . It is noted that the DOA estimation is quite accurate using the superdirective beamformer (0° from Table III), and the discrepancy of the subjective test is due predominantly to the HRTF database. In summary, these results reveal that the developed system is effective in creating a spatial sound field that allows for the practical localization needs of human listeners.

IV. CONCLUSIONS

A microphone array accompanied with 3-D spatial post-processing system has been developed in this paper. Array signal processing algorithms are implemented on a DSP system to accommodate broadband acoustical applications with better SNR and directivity. To further enhance the directivity at the low-frequency range, the superdirective method is exploited in the beamformer design. The entire system consists of a DOA estimation module, a beam steering module, and

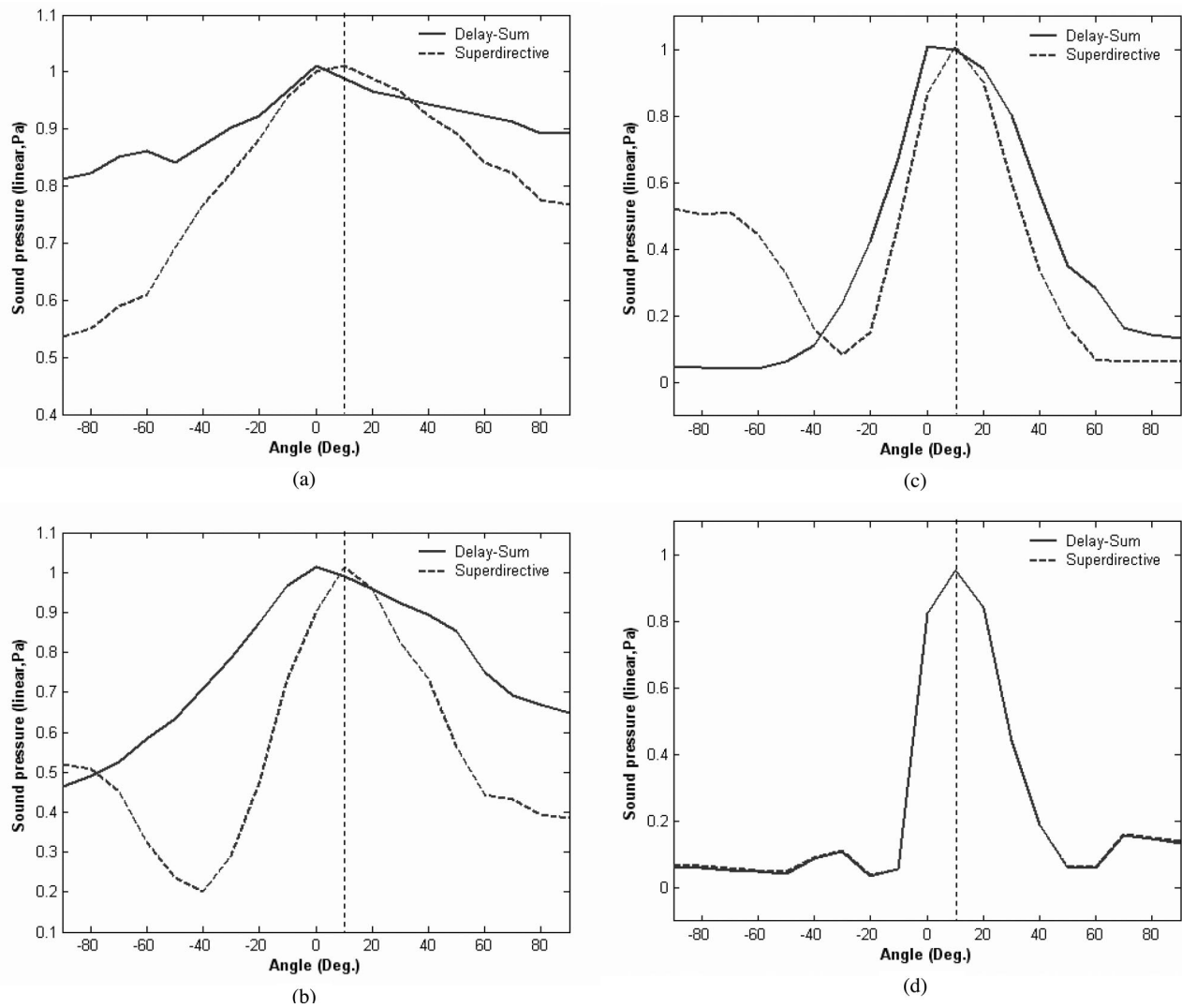


FIG. 10. The experimental results of the DOA estimation (in a linear scale) by the delay-sum and superdirective methods. The sound source is a pure tone and is located at $\theta=10^\circ$, as indicated by a vertical dotted line in the figure. (a) 500 Hz. (b) 1000 Hz. (c) 2000 Hz. (d) 4000 Hz.

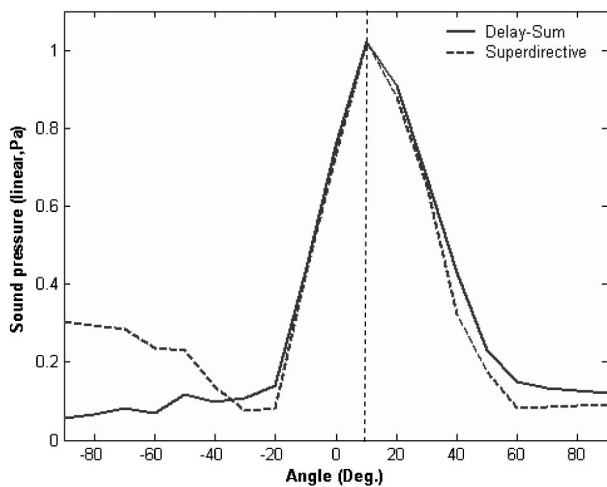


FIG. 11. The experimental results of the DOA estimation by the delay-sum and superdirective methods. The sound source is a random noise bandlimited to 8 kHz and located at $\theta=10^\circ$, as indicated by a vertical dotted line in the figure.

an HRTF post-processing module. It was suggested by the reviewer that, instead of free-field HRTFs, stereo room responses should be used. However, the present paper did not choose this approach because we feel that the natural reverberations and responses of the room environment were not lost in beamforming. What had been lost was the information associated with human head scattering and diffraction. In addition, as another practical reason, filtering of room responses is known to be a computationally expensive operation, which is still prohibitive in the present DSP platform used in the research.

TABLE III. Error analysis of DOA estimation obtained by the delay-sum beamformer and the superdirective beamformer.

Source	Delay-sum beamformer	Superdirective beamformer
500 Hz sine	-10°	0°
1 kHz sine	-10°	0°
2 kHz sine	-10°	0°
4 kHz sine	0°	0°
Broadband random	0°	0°
Average error	6°	0°

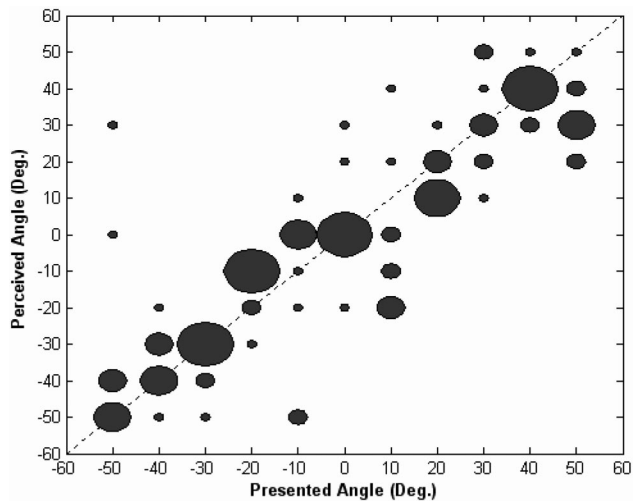


FIG. 12. The results of the subjective listening experiment obtained using the proposed array spatial sound system. The perceived angle (y axis) is plotted versus the presented angle (x axis). The 45° line stands for perfect localization.

Simulations and experiments were carried out to investigate the feasibility of the proposed array system. It was found from the results that the array has achieved significant improvement in terms of SNR as well as directivity. A subjective listening experiment was conducted to examine the source localization performance using the present system. The results indicate that the subjects are capable of localizing the presented source direction within an average error of 11.7°.

Although these preliminary tests reveal the potential of the present system, several limitations of the work should be

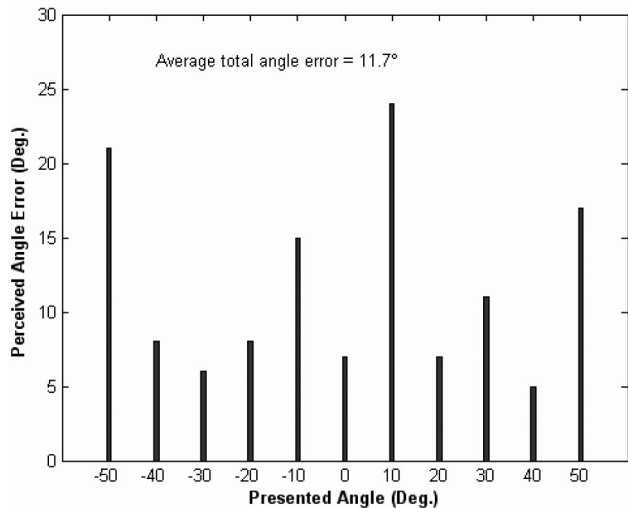


FIG. 13. The bar chart showing the statistics of localization errors of the subjective listening experiment for each presented angle. The average localization error is 11.7° and the standard deviation is 6.5°.

mentioned. First, the selected sampling rate of 16 kHz was somewhat low because of the limited processing power of the present DSP we used. This gives an effective frequency range of less than 8 kHz, which might be insufficient for sound signals other than speech. However, this limitation is minor and can be removed by using a faster DSP. Second, the system works quite well with continuous and stationary signals but may lose track of the source for transient signals. Tracking algorithms suited to transient signals remain to be investigated. Third, background noise and reflection in a live room may seriously interfere the DOA estimation of the present system—a problem common to array systems. It is worth exploring in the future research algorithms for the DOA estimation that are both sharp and robust in a reverberant environment.

ACKNOWLEDGMENT

The work was supported by the National Science Council in Taiwan, Republic of China, under Project No. NSC 92-2212-E009-030.

- ¹P. L. Chu, "Superdirective microphone array for a set-top videoconference system," *Proceedings of the IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, 1997.
- ²D. Korompis, K. Yao, R. E. Hudson, and F. Lorenzelli, "Microphone array signal processing for hearing aid application," *IEEE Workshop on VLSI SP*, Chicago, 1994.
- ³J. M. Kates and M. R. Weiss, "A comparison of hearing-aid array processing techniques," *J. Acoust. Soc. Am.* **99**, 3138–3148 (1996).
- ⁴D. Korompis, A. Wang, and K. Yao, "Comparison of microphone array designs for hearing aid," *Proceedings of the IEEE ICASSP*, 1995.
- ⁵C. Liu and S. Sideman, "Simulation of fixed microphone arrays for directional hearing aids," *J. Acoust. Soc. Am.* **100**, 848–856 (1996).
- ⁶G. H. Saunders and J. M. Kates, "Speech intelligibility enhancement using hearing-aid array processing," *J. Acoust. Soc. Am.* **102**, 1827–1837 (1997).
- ⁷H. L. Van Trees, *Optimum Array Processing* (Wiley, New York, 2002).
- ⁸H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.* **35**, 1365–1375 (1987).
- ⁹H. Cox, R. M. Zeskind, and T. Kooij, "Practical supergain," *IEEE Trans. Acoust., Speech, Signal Process.* **34**, 393–398 (1986).
- ¹⁰J. M. Kates, "Superdirective arrays for hearing aids," *J. Acoust. Soc. Am.* **94**, 1930–1943 (1993).
- ¹¹R. W. Stadler and W. Rabinowitz, "On the potential of fixed arrays for hearing aids," *J. Acoust. Soc. Am.* **94**, 1332–1342 (1993).
- ¹²J. Bitzer, K. U. Simmer, and K. D. Kammeyer, "An alternative implementation of the superdirective beamformer," *Proceedings of the IEEE Workshop on Application of Signal Processing to Audio and Acoustics*, 1999, pp. 17–20.
- ¹³M. Brandstein and D. Ward, *Microphone Arrays: Signal Processing Techniques and Applications* (Springer-Verlag, New York, 2001).
- ¹⁴B. Gardner and K. Martin, "HRTF measurements of a KEMER dummy-head microphone," MIT Media Lab., 1994.
- ¹⁵T. I. Laakso, V. Valimaki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay," *IEEE Signal Processing Magazine*, January, 1996, pp. 30–60.
- ¹⁶O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE* **60**, 926–935 (1972).
- ¹⁷G. W. Elko, "Superdirectional microphone arrays," *Acoustical Signal Processing for Telecommunication*, edited by S. Gay and J. Benesty (Kluwer Academic, Boston, 2000).

The shift-invariant discrete wavelet transform and application to speech waveform analysis

Jörg Enders, Weihua Geng, Peijun Li, and Michael W. Frazier
Department of Mathematics, Michigan State University, East Lansing, Michigan 48824-1027

David J. Scholl
Ford Motor Company, MD3083/SRL Building, Dearborn, Michigan 48121-2053

(Received 17 April 2004; revised 15 October 2004; accepted 24 January 2005)

The discrete wavelet transform may be used as a signal-processing tool for visualization and analysis of nonstationary, time-sampled waveforms. The highly desirable property of shift invariance can be obtained at the cost of a moderate increase in computational complexity, and accepting a least-squares inverse (pseudoinverse) in place of a true inverse. A new algorithm for the pseudoinverse of the shift-invariant transform that is easier to implement in array-oriented scripting languages than existing algorithms is presented together with self-contained proofs. Representing only one of the many and varied potential applications, a recorded speech waveform illustrates the benefits of shift invariance with pseudoinvertibility. Visualization shows the glottal modulation of vowel formants and frication noise, revealing secondary glottal pulses and other waveform irregularities. Additionally, performing sound waveform editing operations (i.e., cutting and pasting sections) on the shift-invariant wavelet representation automatically produces quiet, click-free section boundaries in the resulting sound. The capabilities of this wavelet-domain editing technique are demonstrated by changing the rate of a recorded spoken word. Individual pitch periods are repeated to obtain a half-speed result, and alternate individual pitch periods are removed to obtain a double-speed result. The original pitch and formant frequencies are preserved. In informal listening tests, the results are clear and understandable. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1869732]

PACS numbers: 43.60.Hj, 43.72.Ar [DOS]

Pages: 2122–2133

I. INTRODUCTION

In experimental acoustics, it is common to encounter nonstationary sound waveforms, i.e., those in which the frequency content and amplitude change as a function of time. The conventional approach for analyzing such sounds is to calculate a spectrogram, or short-time Fourier transform (STFT). For a time-sampled waveform $z(t)$, the STFT provides information about the waveform's energy content as a function of both time and frequency, i.e., $F_{\text{STFT}}(z) = E(t, f)$. While the STFT has proven its worth in numerous practical applications, it is ill suited to certain types of sounds, and it lacks some desirable mathematical characteristics. Sounds with frequency content ranging over more than one or two orders of magnitude are often problematic for STFT analysis, because a window long enough to capture low-frequency content (at least one period) will be insensitive to high-frequency sounds of short time duration.

The discrete wavelet transform (DWT) has a severe limitation when used for acoustic waveform analysis: its lack of shift invariance. Let two time-sampled waveforms $z(t)$ and $z'(t)$ be time-shifted copies of one another, such that $z'(t) = z(t + t_0)$ for all t . Calculating the DWT of each, $F_{\text{DWT}}(z) = E(t, n)$, and $F_{\text{DWT}}(z') = E'(t, n)$. Since the DWT is not shift invariant, $E(t, n) \neq E'(t + t_0, n)$. Therefore, the DWT analysis of a sampled sound depends on when the sampling starts, not just when the sound occurs, which is highly undesirable for the study of physical systems. The DWT is critically sampled, i.e., utilizes lower sampling rates

(subsampling) for lower-frequency components. The selection of samples to be skipped in the subsampling process is inextricably linked to the time elapsed since the sampling began. Fortunately, by modifying the DWT to retain all possible samples (performing no subsampling), it is possible to obtain explicit shift invariance.¹⁻³ The resulting shift-invariant discrete wavelet transform (SIDWT) is highly redundant, but since many of the redundant elements are duplicates, the increase in computational complexity is not severe. The full SIDWT may be used as a starting point from which to draw a more efficient representation for lossy compression.⁴ The SIDWT may also be used in full (albeit with the duplicate elements grouped and summed), in which form it has been shown to be an isometry, with applications in data visualization.⁵ Others have described algorithms which are mathematically equivalent to the SIDWT, but which were developed for applications in exploratory statistics, using different nomenclature, i.e., the stationary wavelet transform⁶ and the maximal overlap discrete wavelet transform.⁷ The stationary wavelet transform has also been used successfully for waveform denoising.²

The use of the SIDWT (and its equivalents) to identify features in a waveform; whether signatures of interesting phenomena, experimental artifacts, or noise, leads naturally to the following question. What would the time-sampled waveform look like (or sound like) if the features were louder, softer, appeared at a different time, or were removed altogether? Performing the desired modifications on the SIDWT output is straightforward; the challenge is reversing

the SIDWT to return to a time-sampled waveform. Because of its inherent redundancy, the SIDWT does not have a true inverse in the mathematical sense. However, this fact does not preclude the existence of an algorithm with useful inverse-like behavior. The developers of the stationary wavelet transform also developed such an inverse-like procedure. They showed that averaging together all of the possible shift-induced variations of the IDWT yields intuitively satisfying results.⁶ Likewise, a mathematically equivalent procedure was used to invert the maximal overlap discrete wavelet transform, also with intuitively satisfying results.⁷ The stationary wavelet transform combined with this inverse-like averaging procedure has also been shown to yield good results in waveform denoising.² The wavelet denoising paper states, without proof or discussion, the important mathematical result that the inverse-like procedure is actually the pseudoinverse of the stationary wavelet transform. A recent publication coauthored by one of the present authors describes two examples of sound visualization and modification using the SIDWT and its pseudoinverse (ISIDWT). The discussion and the two examples are narrowly focused on the field of automotive sound quality engineering, and no mathematical material is included.⁸

The goals of this paper are threefold. The first is to describe a newly developed simple and fast convolution algorithm for the ISIDWT, based on the SIDWT algorithm.⁵ The SIDWT, the stationary wavelet transform, and the maximal overlap discrete wavelet transform employ significantly different algorithms, so a discussion of computational issues is included. The second goal is to present a simple, self-contained proof that the ISIDWT is the pseudoinverse of the SIDWT. The statement of this result has been published; we believe the details of the proof should be made available as well. The third goal is to illustrate the potential applications of these new analytical methods in the field of acoustics. Section II covers both the theoretical (II A) and the computational (II B) aspects of the SIDWT and its pseudoinverse. In Sec. III, examples of low-level speech waveform processing illustrate the capabilities of the SIDWT/ISIDWT for visualization, feature separation, and analysis/synthesis. One especially promising way to combine these capabilities is to edit (cut and paste sections) of sound recordings in the shift-invariant wavelet domain. While many audible features are easier to recognize in that domain, the primary benefit is the wavelet pseudoinverse transform automatically prevents the occurrence of the audible clicks and pops that are usually produced at section boundaries by time-domain editing. Illustrative examples of waveforms with strong time localization and a wide frequency range can be found in many different technical fields of study. For the development of digital audio effects in music, it is useful to be able to distinguish transient (time-localized) sounds, such as the pluck of a guitar string, from the steady ringing tone (frequency-localized) that follows.⁹ Research in wavelet-domain modification of musical sounds began in the early days of wavelet theory, e.g., musical applications of complex wavelets,¹⁰ and continues today, e.g., the use of a “lapped” wavelet transform to stitch together segments of musical waveforms.¹¹ The examples most familiar to the authors are drawn from

the myriad of mechanical sounds produced by motor vehicles, e.g., a momentary rattle excited by (and partially masked by) a car door slam, and the motor whine, blade scrape, and reversal thud of a windshield wiper.⁸ The details of speech waveforms, especially the formant resonances modulated by glottal pulses, are also an excellent match to the capabilities of the SIDWT/ISIDWT. The general approach and terminology derives from an *ad hoc* list of recent publications that deal with various details of speech waveforms: pitch period estimation,^{12–14} formant modulation,¹⁵ friction noise modulation,¹⁶ voicing onset,¹⁷ glottal characteristics,¹⁸ and waveform irregularities.¹⁹

II. THE SHIFT-INVARIANT DISCRETE WAVELET TRANSFORM AND ITS PSEUDOINVERSE

A. Theory

Consider a sequence of N physical measurements $z = (z_1, z_2, \dots, z_N)$, e.g., air pressure measured repeatedly at evenly spaced time intervals. Let S denote the set of all such signals. Since $\sum_n z_n^2 < +\infty$, the vector z may be regarded as the coordinates of a single point in a finite energy, N -dimensional vector space, $z \in l^2(\mathbb{Z}_N)$. Implicit in $z \in l^2(\mathbb{Z}_N)$ is the assumption that z is a single period of an infinitely long sequence with a periodicity of N . If this assumption is not physically realistic, care must be taken to insure that the conclusions drawn from the analysis are independent of N .

Let $u \in l^2(\mathbb{Z}_N)$ and $v \in l^2(\mathbb{Z}_N)$ represent two digital filters. Denoting the discrete Fourier transform of u by \hat{u} , we require the system matrix

$$A(n) = \frac{1}{\sqrt{2}} \begin{pmatrix} \hat{u}(n) & \hat{v}(n) \\ \hat{u}\left(n + \frac{N}{2}\right) & \hat{v}\left(n + \frac{N}{2}\right) \end{pmatrix}, \quad (1)$$

to be unitary for each $n = 0, \dots, N-1$ (Ref. 20, p. 173). Therefore, u is the low-pass filter sequence and v the high-pass filter sequence generating the discrete wavelet transform.

Let \bar{u} be the complex conjugate reflection of u defined by $\bar{u}(n) = u^*(N-n)$ for all n . The finite impulse response filtering of z by u is written as a (circular) convolution $z * u$. Most practical applications of these techniques, including the examples presented here, involve only N -element sequences of real numbers, i.e., $x \in \mathbb{R}^N \subset l^2(\mathbb{Z}_N)$. The mathematical results, however, are valid for complex-valued vectors. Assuming that m divides N , a sequence reordering operator R_m , defined by

$$R_m(z) = (z_1, z_{m+1}, \dots, z_{N-(m-1)}, z_2, z_{m+2}, \dots, z_{N-(m-2)}, \dots, z_m, z_{2m}, \dots, z_N), \quad (2)$$

in effect, writes the elements of z into an m by N/m matrix by columns and reads the elements out by rows. For example, if $z = (1, 2, 3, 4, 5, 6, 7, 8)$, then $R_2(z) = (1, 3, 5, 7, 2, 4, 6, 8)$. The inverse of R_m is $R_{N/m}$, i.e., $R_{N/m}(R_m(z)) = z$.

From an N -element input vector, given p such that 2^p divides N , the p -stage shift-invariant discrete wavelet trans-

form T produces a $(p + 1)$ by N matrix.⁵ T may therefore be regarded as a linear map taking each point in S to a point in a $(p + 1)N$ -dimensional vector space \mathbf{W} , i.e.,

$$T: l^2(\mathbb{Z}_N) \rightarrow l^2(\mathbb{Z}_{(p+1)N}). \quad (3)$$

T of z is given by

$$T(z) = (R_{N/2}(x_1), R_{N/4}(x_2), \dots, R_{N/2^p}(x_p), R_{N/2^p}(y_p)), \quad (4)$$

where

$$x_1 = \frac{1}{\sqrt{2}} R_2(z * \tilde{v}) \quad \text{and} \quad y_1 = \frac{1}{\sqrt{2}} R_2(z * \tilde{u}), \quad (5)$$

and for $j = 2, \dots, p$

$$x_j = \frac{1}{\sqrt{2}} R_2(y_{j-1} * \tilde{v}) \quad \text{and} \quad y_j = \frac{1}{\sqrt{2}} R_2(y_{j-1} * \tilde{u}). \quad (6)$$

The set of points mapped by T from S occupies a subspace in \mathbf{W} denoted by range (T).

Being a linear map from an N -dimensional space of signals to a space of larger dimension, the SIDWT does not have an inverse. Of all the points in \mathbf{W} , only those which are in range (T) are directly associated with a point in S . The pseudoinverse works around this limitation by providing every point $w \in \mathbf{W}$ with an *indirect* association to some point $z \in S$. Every w has a unique nearest (in the standard Euclidean norm) neighbor $w' \in \text{range}(T)$ (possibly itself), and the pseudoinverse associates each w with the z that satisfies $T(z) = w'$. This procedure is mathematically equivalent to finding the least-squares solution to an overdetermined system of linear equations. We now define the ISIDWT

$$S: l^2(\mathbb{Z}_{(p+1)N}) \rightarrow l^2(\mathbb{Z}_N), \quad (7)$$

a map taking each point in \mathbf{W} to a point in S . Given $w = (w_1, w_2, \dots, w_{p+1}) \in l^2(\mathbb{Z}_{(p+1)N})$, we compute $S(w)$ by the algorithm

$$\eta_p = \frac{1}{\sqrt{2}} (R_{N/2}(R_{2^p}(w_p)) * v + R_{N/2}(R_{2^p}(w_{p+1})) * u) \quad (8)$$

$$\eta_{p-1} = \frac{1}{\sqrt{2}} (R_{N/2}(R_{2^{p-1}}(w_{p-1})) * v + R_{N/2}(\eta_p) * u) \quad (9)$$

\vdots

$$S(w) = \eta_1 = \frac{1}{\sqrt{2}} (R_{N/2}(R_2(w_1)) * v + R_{N/2}(\eta_2) * u). \quad (10)$$

The relationship between T and S (as defined above) is established by the following theorem.

Theorem 1. S is the least-squares inverse (pseudoinverse) of T , i.e.,

- (i) $ST = \text{id}|_{l^2(\mathbb{Z}_N)}$, and
- (ii) TS is the orthogonal projection of $l^2(\mathbb{Z}_{(p+1)N})$ onto range (T).

According to statement (i), for $w \in \text{range}(T)$, S is the inverse of T , and therefore $S(T(z)) = z$. Statement (ii) ad-

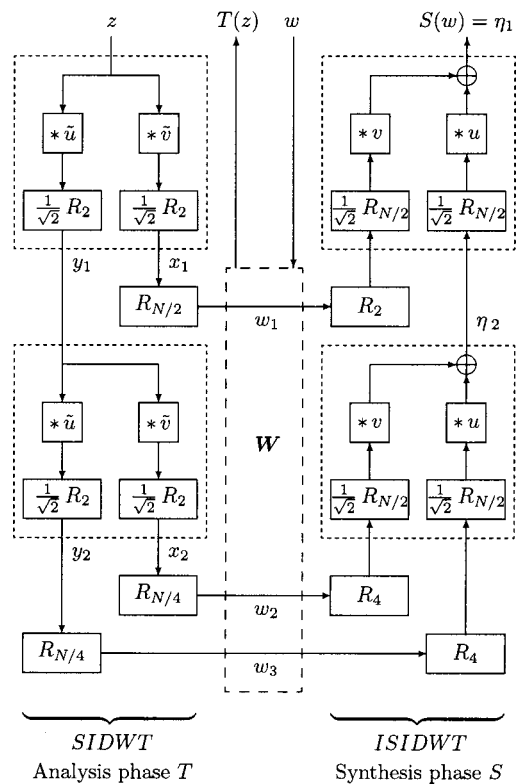


FIG. 1. A block diagram of the SIDWT and ISIDWT for two scale levels, i.e., $p = 2$.

resses the case of $w \notin \text{range}(T)$, which will apply to virtually all w chosen arbitrarily, i.e., not obtained by $w = T(z)$. In this case, $x = T(S(w)) \in \text{range}(T)$ is the unique point in range (T) closest to w , minimizing $\sum_n (x_n - w_n)^2$. A proof of Theorem 1 is given in the Appendix.

B. Computation

The block diagram in Fig. 1 illustrates the processing steps in the SIDWT (analysis phase T) and its pseudoinverse ISIDWT (synthesis phase S) for the two-stage case, i.e., $p = 2$. The analysis begins in the upper-left corner with the input waveform z . The four-block clusters inside the dotted-line boxes on the left side of the diagram depict the recursive analysis steps defined in Eqs. (5) and (6). A p -stage transform employs p of these clusters, yielding $p + 1$ many series $x_1, x_2, \dots, x_p, y_p$, each of length N . The final rearrangement step defined in Eq. (4) reaches the vector space \mathbf{W} , depicted by the central dashed-line box. Completion of the SIDWT is indicated by the vertical arrow leading to $T(z)$ at the top of the diagram. The synthesis begins at the top with w , which in most cases will be a modification of $T(z)$. The first step in the ISIDWT is to undo the final rearrangement step in the SIDWT. After this, the recursive procedure in the dotted-line boxes in the right side of the diagram, as defined in Eqs. (8) and (9), is carried out. The ISIDWT is complete at the top of the diagram where $S(w) = \eta_1$, as in Eq. (10).

Let us look at the computational complexity of the pseudoinverse S . Given $z \in l^2(\mathbb{Z}_N)$, the computational complexity for the transform $T(z)$ is $\mathcal{O}(N \log_2 N)$ according to Ref. 5. We will now show that this result holds for the ISIDWT S as well.

Theorem 2. Let N be a power of 2 and $p \in \mathbb{N}$, with $p \leq \log_2 N$ fixed. Then, the total number of complex multiplications required to compute $S(w)$ for $w \in l^2(\mathbb{Z}_{(p+1)N})$ is $\#_N \leq 2pN + 3pN \log_2 N$.

Proof: In lieu of a detailed proof, we note that S and the analysis algorithm are essentially the same by symmetry.

From Theorem 2 we obtain the result that the computation of S is an $\mathcal{O}(N \log_2 N)$ operation if p is considered to be a fixed number. If we take the Daubechies D4 u and v and perform the convolutions directly instead of using the FFT, then the computational complexity is just $\mathcal{O}(N)$, since u and v have only 4 nonzero entries. This is the minimal order we can expect when working with signals of length N . Since S is a linear map, the error in the output is bounded by the norm of S (a constant) times the error in the input, which means that the algorithm is numerically stable.

The ISIDWT, the stationary wavelet transform, and the maximal overlap discrete wavelet transform are mathematically equivalent in the sense that they yield the same result. However, they utilize significantly different algorithms, so they are not computationally equivalent in all respects. All three may be calculated with $\mathcal{O}(N \log_2 N)$ computational complexity if p is considered fixed.⁵⁻⁷ If implemented in a low-level programming language that allows efficient indexing of individual matrix elements, the performance of the three is expected to be essentially equivalent. However, the ISIDWT is significantly easier to implement in an array-oriented scripting language, because it can be constructed by linking together a few of the standard functions that are commonly provided in such languages. In this way, acceptable performance can be obtained without the need to write, compile, and link an external module written in a lower-level language.

III. APPLICATION TO SPEECH WAVEFORM ANALYSIS

To illustrate the application of the SIDWT/ISIDWT to acoustic waveforms, a detailed analysis of a sound recording of a spoken word is presented below. A recording of a male speaker pronouncing the Japanese word “kaze” with a rising intonation from an on-line speech database maintained for phonetic alphabet research²¹ is shown in Fig. 2. Voiced speech is produced by periodic glottal closure events, which momentarily interrupt the air flow through the larynx. The frequency at which these events occur, denoted by F_0 , is the fundamental frequency of voiced speech. The rising intonation in this example is reflected in Fig. 2, as $F_0 \approx 90$ Hz during the “a” increases to $F_0 \approx 120$ Hz during the “e.” The procedure for glottal period estimation is discussed in detail below.

During spoken vowels, the sharp air-pressure transients known as glottal pulses excite pressure oscillations in the volume acoustic resonances of the vocal tract. The frequency content of these resonances, which fall in the range between ~ 500 Hz to ~ 8 kHz, depending on the size of the vocal tract and the position of the tongue, jaw, and lips, is the primary factor distinguishing one vowel from another. The significant frequency peaks in the pressure oscillations are known as formants, and are denoted F_1, F_2, F_3, F_4 , in order of increasing frequency. The formant amplitude is highest immediately

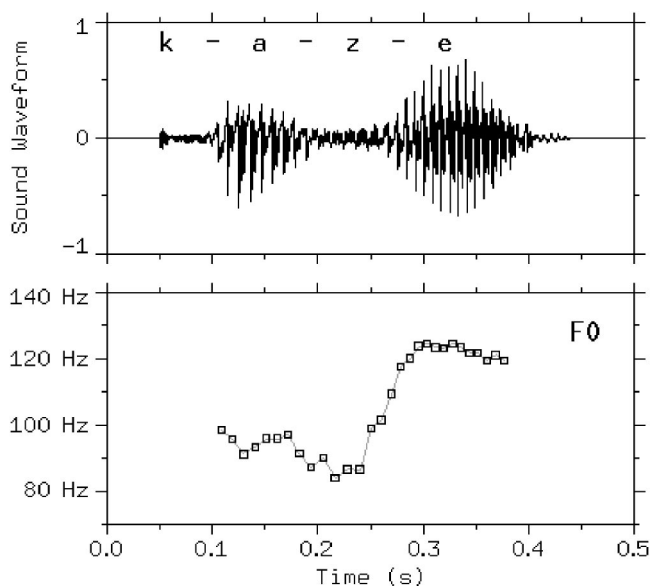


FIG. 2. A sound recording of a male speaker pronouncing the Japanese word “kaze” with rising intonation, and the voiced fundamental frequency F_0 obtained from glottal period estimates.

after the glottal pulse, at which point the glottis is essentially closed. The amplitude decreases rapidly as energy is lost due to air flow between the lips. When the glottis reopens in preparation for the next glottal pulse, the resulting air flow causes the formant amplitude to decrease even more quickly. This strong amplitude modulation of the formants leads to a widely used quasistatic approximation. In this simplified picture, the frequency content of each formant pulse is assumed to be static, and the (relatively slow) motion of the vocal-tract anatomy (tongue, jaw, and lips) is inferred by comparing the frequency content of consecutive formant pulses. The first six formant pulses in kaze may be seen in Fig. 3. The first formant pulse, which signifies the beginning of the vowel sound “a,” occurs at ~ 0.105 ms. In the sound waveform plot, each formant peak begins at a sharp downward step (a glottal pulse) and oscillates with decreasing amplitude, disappearing before the next glottal pulse. On the spectrogram, labeled “STFT,” each downward step appears as a vertical gray bar; the short time duration of each step maps to broad frequency content. The sampling rate was 44 100 samples/s, and a 352-point Hanning window, shifted in 63-point steps, and zero-padded to 1024-point length, was used in the preparation of spectrograms in this and the next two figures. The gray scale on each spectrogram was adjusted for enhanced contrast.²² Each formant pulse appears as a pair of dark horizontal bands ~ 600 Hz apart, beginning at a vertical bar, and ending before the next vertical bar. This formant frequency content of $F_1 \approx 500$ Hz and $F_2 \approx 1100$ Hz is typical for a male Japanese speaker’s “a.”²³

The scalogram in Fig. 3 labeled “DWT” is obtained from the conventional, shift-variant, discrete wavelet transform. The 8-tap symlet was used for all examples presented here, but the results do not depend critically on the choice of wavelet. The shape of the wavelet (e.g., symlet vs Daubechies) makes little difference here. Shorter wavelet filters (e.g., 4-tap vs 8-tap) will have increased energy in the

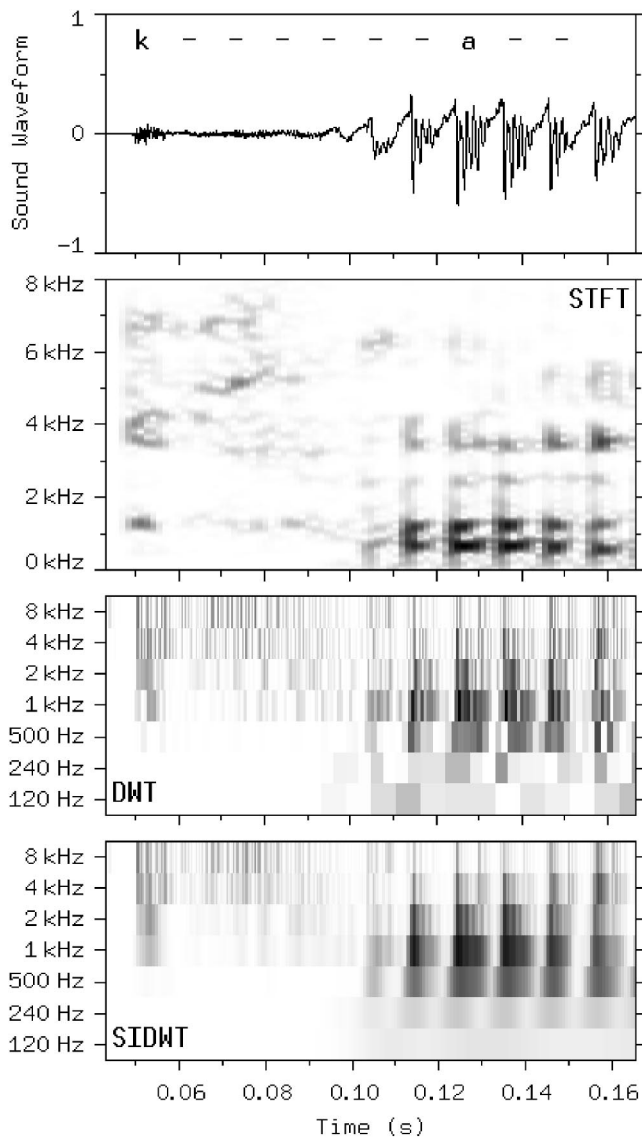


FIG. 3. The “ka” from the Japanese word “kaze,” its spectrogram (STFT), its conventional scalogram (DWT), and its shift-invariant scalogram (SIDWT).

upper sidebands, but provide faster calculations and sharper time resolution. The essential preprocessing step for visualization is to transform the oscillatory coefficients within each scale level of $w \in \text{range}(T)$ to a quadratic envelope.⁵ For the coefficients at the m th scale level, w_m , the quadratic envelope $w'_m = w_m^2 + H(w_m)^2$, where H is the Hilbert transform, i.e., a $\pi/2$ phase shift. Following this operation, all of the scalograms presented here were downsampled to fit the available space, and the gray scales were adjusted for enhanced contrast.²² The formant pulses appear as dark, vertical features extending from the 500-Hz to the 8-kHz bands. Their appearance is more varied than on the spectrogram, due to the shift variance of the DWT. Nevertheless, the DWT has been shown to be a reliable method for identifying glottal pulses for F_0 estimation.^{13,14} The scalogram in Fig. 3 labeled “SIDWT” is obtained from the shift-invariant discrete wavelet transform. In the region corresponding to the “k” sound, the SIDWT and DWT scalograms have a similar appearance. In the region corresponding to the voiced “a,”

they are dramatically different. The shift invariance reveals the true reproducibility of the formant pulses, and for each formant pulse shows the amplitude decrease and the gap preceding the next glottal pulse. The voiced region highlights two important differences between the STFT presentation and the SIDWT presentation. First, the STFT has finer frequency resolution than the SIDWT. The two main formants, F_1 and F_2 , are resolved clearly on the STFT, but are not resolved on the SIDWT due to the single-octave bandwidth of the scalogram levels. The second difference is that the STFT has the same time resolution at all frequencies, so transients appear as vertical features. In contrast, the time resolution of the SIDWT scales inversely with the center frequency of each band. For each step upward to a higher-frequency scalogram band, the time resolution is twice as fine. For this reason, a transient feature tends to have a pyramidal appearance on the SIDWT, with a narrow top on a base that broadens at each next lower level. The practical consequence of these two differences is that the SIDWT is not a substitute or a replacement for the STFT, but rather a complement, and the two techniques can be used effectively together.

Figure 4 shows a similar presentation of the “z” from kaze. This sound is produced by a narrow restriction in the mouth. The frictional (turbulent) loss due to the air flowing through the restriction prevents the build-up of formants. The absence of formants does not imply silence, however, because the turbulence produces audible noise called frication. The loudness of the frication varies with the flow of air through the restriction, which in turn is modulated by periodic glottal closures. The modulated frication appears as bursts of noise (glottis open) separated by momentary silences (glottis closed).¹⁶ It is interesting to contrast this timing to that observed with formants, which are loudest when the glottis is closed, and quiet when the glottis is open.

To complete the presentation of kaze, the final vowel “e” is shown in Fig. 5. The time-domain clarity of the formant peaks in the shift-invariant scalogram (SIDWT), compared to the shift-variant scalogram (DWT), is even more evident here than in Fig. 3. The formant pulses are closer together than in Fig. 3, and they also exhibit a second high-frequency pulse in each glottal period. Secondary glottal pulses such as these are often observed in male speech waveforms, and they can be problematic for glottal period estimation algorithms. Interestingly, the phenomenon is usually vowel-dependent, and only traces of secondary pulses can be seen on the “a” in Fig. 3.

An expanded view of two of the formant pulses from Fig. 3 is shown in Fig. 6, as a time history and as a shift-invariant scalogram. The glottal pulses are indicated by “GP.” A periodic signature with a period of ~ 1.7 ms is apparent on the 2-kHz scale as alternating bands of light and dark, and is barely visible on the 1-kHz band. The scalogram as prepared for display is a quadratic function of the wavelet coefficients, and this signature is the difference frequency between the two formant peaks, $1.7 \text{ ms} \approx 1/(F_2 - F_1)$. The difference frequency shows up most clearly on the 2-kHz band because there is significant frequency overlap between adjacent bands, and the strong fundamental F_0 and its strong

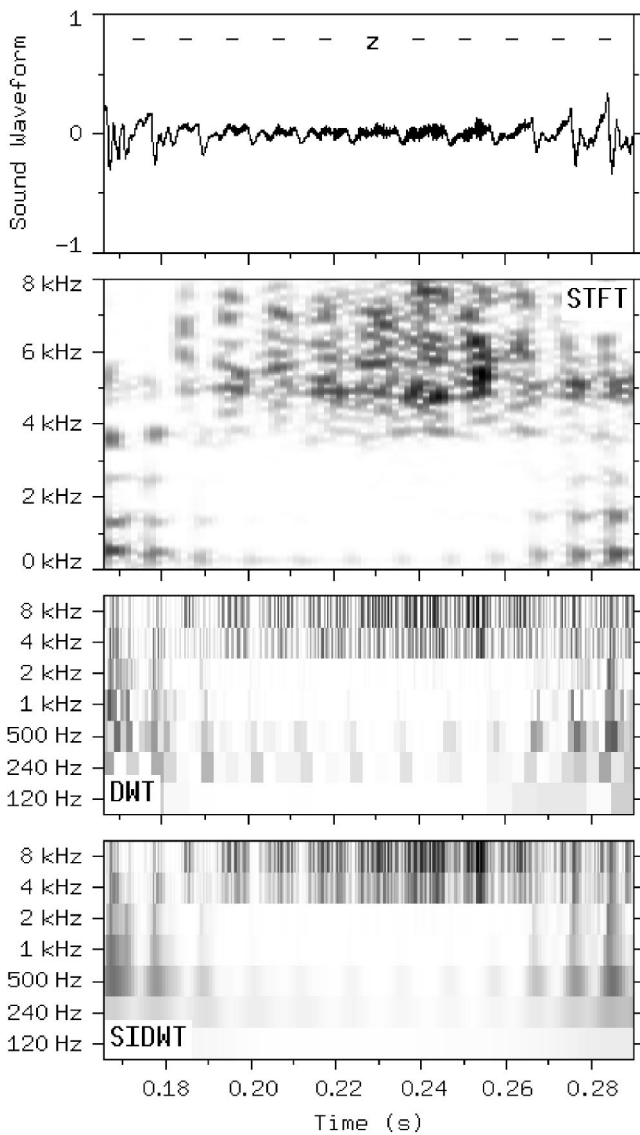


FIG. 4. The “z” from the Japanese word “kaze,” its spectrogram (STFT), its conventional scalogram (DWT), and its shift-invariant scalogram (SIDWT).

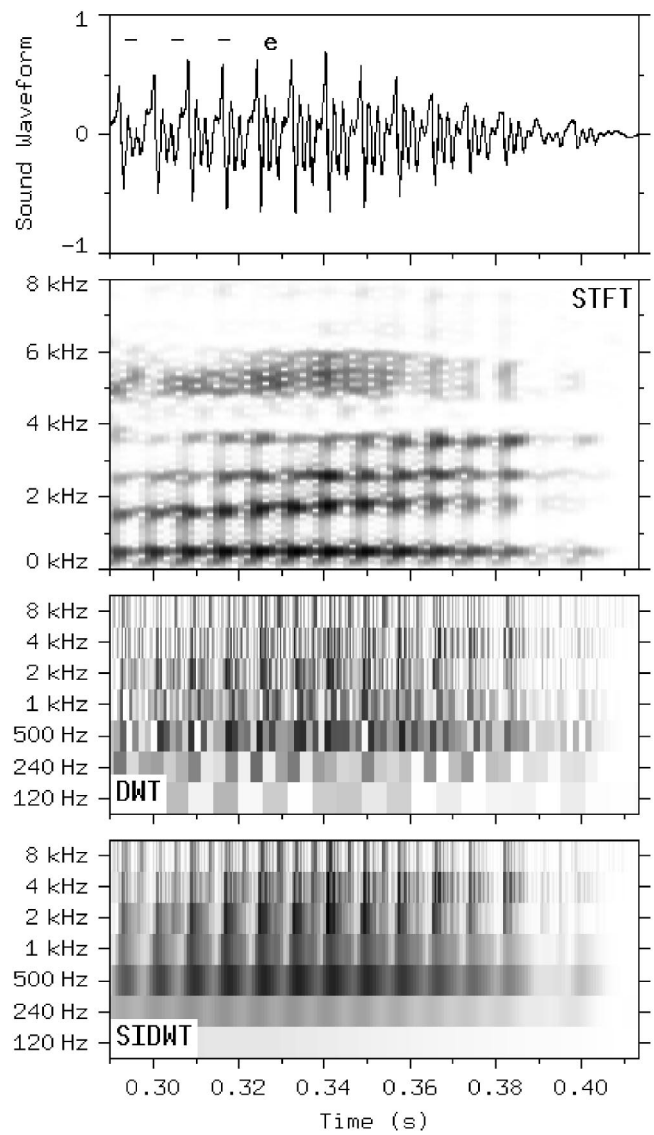


FIG. 5. The “e” from the Japanese word “kaze,” its spectrogram (STFT), its conventional scalogram (DWT), and its shift-invariant scalogram (SIDWT).

second harmonic overwhelm the $F2 - F1$ difference in the 500-Hz and 1-kHz bands.

A similar view of three of the formant pulses from the vowel sound “e” is shown in Fig. 7, as a time history and as a shift-invariant scalogram. Patterns that appear to be difference frequencies can be seen, but since the formant content of “e” is more complex than the two strong peaks responsible for the signature in Fig. 6, the scalogram signature of the vowel “e” is more complex as well. This expanded view provides a more detailed picture of the secondary glottal pulse labeled “2” in each glottal period, showing that the fundamental periodicity, as well as the gap preceding the glottal closure, are still evident. The F_0 estimate in Fig. 2 was obtained by finding all occurrences of this formant gap-peak signature. For a list of times $t = (t_1, t_2, \dots, t_M)$ at which the M occurrences of the signatures were observed, F_0 at t_n is given by $F_{0n} = 1/(t_{n+1} - t_n)$, where $N = 1, 2, \dots, (M - 1)$. The times t were obtained by finding local maxima in the sum of the quadratic envelopes of the 1- and 2-kHz bands.

An important category of speech-processing techniques known as PSOLA (pitch-synchronized overLap and add) is based on working with the individual glottal pulses.²⁴ A typical application of PSOLA might begin with isolating each glottal pulse by multiplying the speech waveform by a rounded window (e.g., Hanning) centered over each pulse in turn. A typical length for the window would be twice the glottal period. The window length represents a compromise: longer windows allow the neighboring pulses to intrude, and shorter windows (or windows with more steeply sloped time-domain cutoffs) increase spectral leakage. After the individual glottal pulses have been processed in the desired manner, they must be recombined to make a single waveform. A variety of approaches has been used to recombine the individual segments,²⁴ including a technique which utilizes information obtained from wavelet transform analysis.²⁵ In general, PSOLA produces high-quality results, although sometimes annoying artifacts are present.^{26,27} The artifacts are not completely understood, and may be related to the

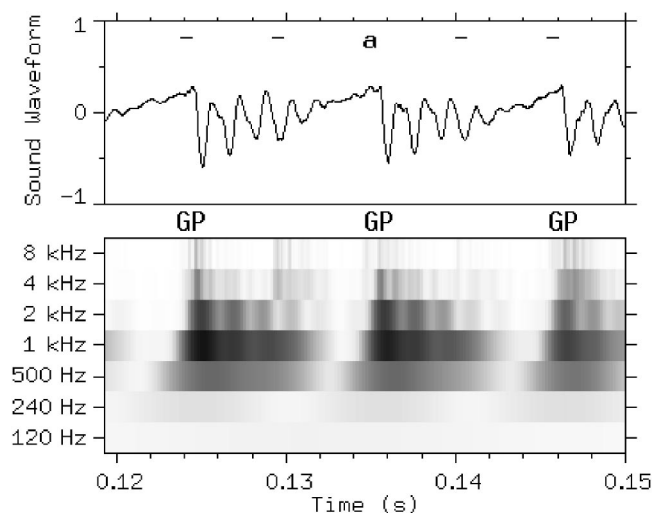


FIG. 6. The Japanese vowel “a,” and its shift-invariant scalogram. The glottal pulses are indicated by “GP.”

details of how the modified pulses are recombined.

The analysis/synthesis capability of the SIDWT/ISIDWT may be employed to segment and recombine a speech waveform in a manner that is conceptually similar to PSOLA, although the mathematical details are of course quite different. Figures 8 and 9 illustrate the procedure for extracting a single formant pulse from the speech waveform. In Fig. 8 a dashed-line box delineates the region of the scalogram corresponding to the single formant pulse to be extracted, i.e., the time interval $0.133 = t_a < t < t_b = 0.1436$. The edges of the box are aligned with local minima of the sum of the quadratic envelopes of the 2- and 1-kHz bands. The scalogram elements inside this box are preserved, and the remainder of the scalogram is set to zero. Given the speech waveform $z(t)$ over the interval $t_0 \leq t \leq t_c$, and the scalogram $x(t, n) = T(z)$, a function of both time and scale, this modification produces $w(t, n)$ such that

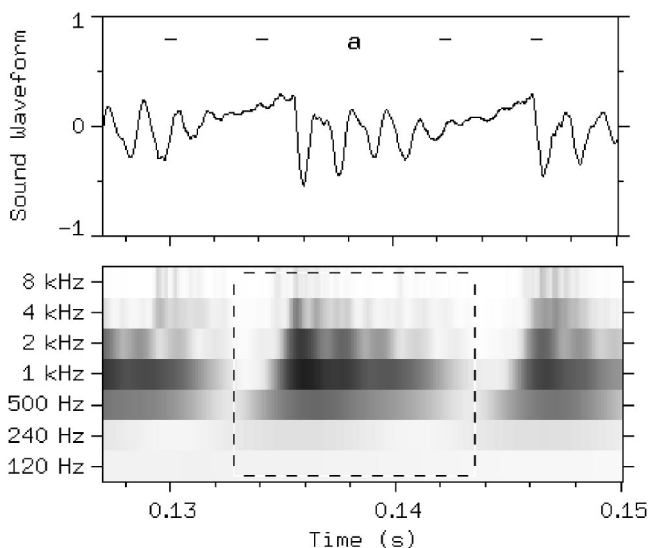


FIG. 8. A single formant pulse signature of the Japanese vowel “a” surrounded by a dashed-line box on the shift-invariant scalogram.

$$w(t, n) = \begin{cases} 0 & : t_0 \leq t < t_a \\ x(t, n) & : t_a \leq t \leq t_b \\ 0 & : t_b < t \leq t_c. \end{cases} \quad (11)$$

To complete the procedure, the extracted single formant pulse z' is obtained from $z' = S(x)$. Figure 9 shows z' , with the original waveform z in gray for comparison. The scalogram $x' = T(z')$ is also shown in Fig. 9, along with the original dashed-line box. The only significant difference between w and x' is the smoothing of the boundaries at t_a and t_b .

To carry the feature extraction procedure described above to completion, the nonzero elements of a scalogram x are segmented into M pieces (x'_1, x'_2, \dots, x'_M) such that $\sum_n x'_n = x$. Then, by the linearity of S , $\sum_n z'_n = z$. The regions of the scalogram where features with strong time localization are absent are segmented at arbitrary time boundaries, with

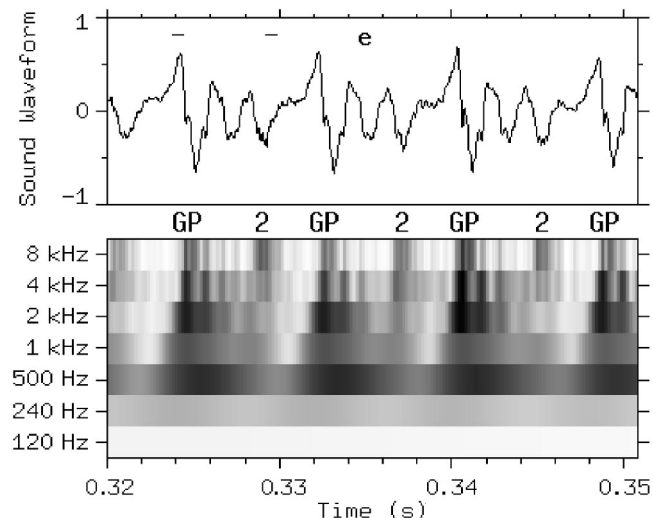


FIG. 7. The Japanese vowel “e,” and its shift-invariant scalogram. The glottal pulses are indicated by “GP,” with secondary pulses indicated by “2.”

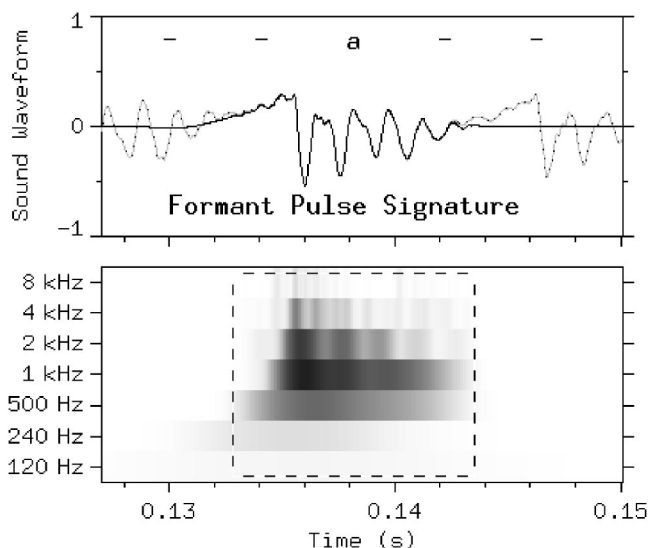


FIG. 9. A single formant pulse of the Japanese vowel “a” extracted by the ISIDWT, with the original waveform in gray for comparison. The shift-invariant scalogram of the reconstructed pulse is shown surrounded by a dashed-line box marking the extracted area.

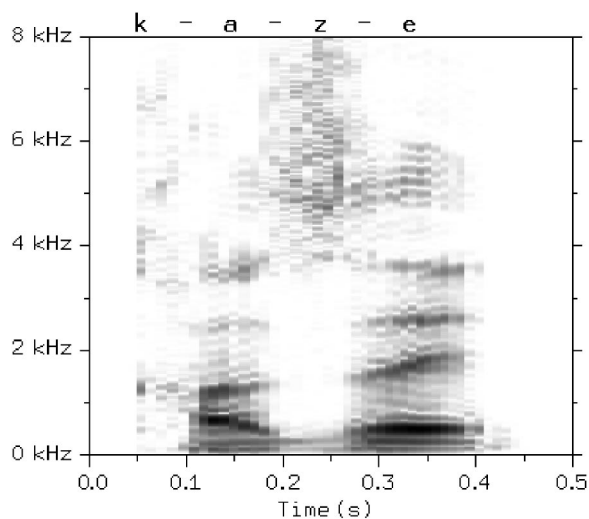


FIG. 10. A pitch-synchronous spectrogram-like display of the Japanese word “kaze.”

the constraint that the arbitrary segmentation lengths $t_b - t_a$ are similar to those used elsewhere in the scalogram. If the segmentation boundaries are aligned with instants of relative quiet, the operation may be considered a type of synchronous windowing, i.e., windowing synchronized with the amplitude modulation inherent in the waveform. In this example, most of the waveform exhibits strong amplitude modulation, and the width of each window $t_b - t_a$ is large relative to the smoothing at the boundaries observed in Fig. 9. Therefore, the time-domain overlap between adjacent segments is negligible, and

$$\sum_{n=1}^M F(z'_n(t)) \approx F(z'(t)), \quad (12)$$

even for some nonlinear $F(z)$ that are sufficiently local, e.g., quadratic envelope, or spectral density for frequencies $f > 1/(t_b - t_a)$.

To show how the frequency content of the extracted pulses evolves over time, a spectrogram-like display is presented in Fig. 10. The formant pulses obtained from the ISIDWT were zero padded to 1536-point length, and the energy spectral density of each pulse was calculated via the FFT with no further windowing. For comparison, a conventional spectrogram of the original kaze waveform is shown in Fig. 11. This and subsequent spectrograms were prepared with a 512-point Hanning window shifted in 134-point steps, and each windowed segment was zero padded to 1536-point length before calculating the FFT.

This waveform segmentation procedure is a unique and powerful capability of the SIDWT/ISIDWT. In addition to the analysis methods shown above, it has broad utility for copying, cutting, and pasting sections of sound waveforms. Working with the scalogram $x = T(z)$ rather than the sound waveform z has two advantages. First, for all but the simplest waveforms, it is usually easier to find and delineate features of interest in x . Second, cutting segments from x and joining them to make w doesn't result in audible clicks and pops. In conventional waveform editing, such clicks and pops are caused by steps at boundaries where the final value in one

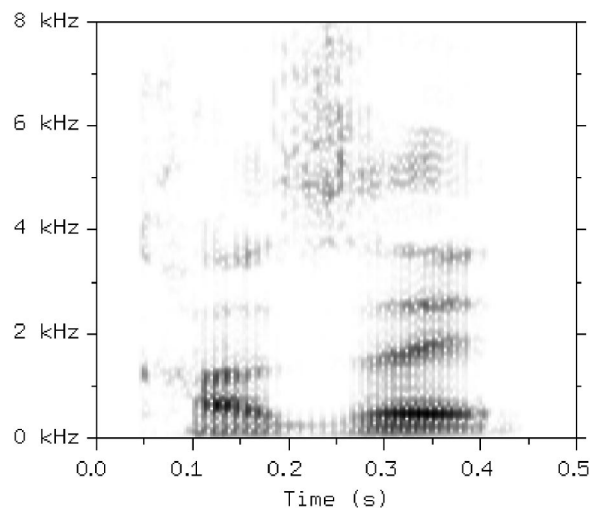


FIG. 11. The spectrogram of the Japanese word “kaze.”

waveform segment differs from the initial value in the next waveform segment. The signature of such a boundary on the scalogram is a peak in the high-frequency scales, reflecting the high-frequency content of the step. It is usually necessary to taper or otherwise reshape the waveforms at each boundary to smooth these steps. However, excessive tapering or reshaping can create other audible artifacts, e.g., a gap in the high-frequency components. In some cases human intervention is required to find the optimal balance. Performing the cutting and joining operations on w creates steps in the scalogram values at boundaries that resemble the steps created on a raw edited sound waveform. However, since they are steps in w , not steps in z , they are not associated with audible clicks and pops. The pseudoinverse smooths the waveform at the boundary in a way that the scalogram signature of the boundary on x' is a rounded step, as close as possible in the least-squares sense to the original sharp step. The pseudoinverse cannot create a click or a pop at the boundary, because that would require a peak on the scalogram x' . A peak at the step location implies scalogram values with magnitude greater than those on either side of the step, which would never be the solution that minimizes $\sum_n (x'_n - x_n)^2$. It should be noted that absence of editing artifacts is no guarantee of realism, since abrupt starts or transitions between sounds of different character may sound false or even unpleasant. Even so, the reliable and automatic prevention of editing artifacts in sound waveforms is a substantial convenience.

A common application of PSOLA is changing the rate of a spoken word without changing the pitch or the frequency content of the formants, i.e., to simulate the same speaker pronouncing the same word, but speaking more rapidly or more slowly. To illustrate the sound waveform editing capabilities of the SIDWT/ISIDWT, the rate of the example waveform kaze has been halved and doubled. The first step in the procedure is to synchronize the analysis with the glottal pulses whenever possible. The spectrogram-like display shown in Fig. 10 was created by identifying M time instants t_i . The t_i in the voiced regions were located at moments of relative quiet in the 1- and 2-kHz bands, and the t_i in the unvoiced regions were merely spaced at regular intervals.

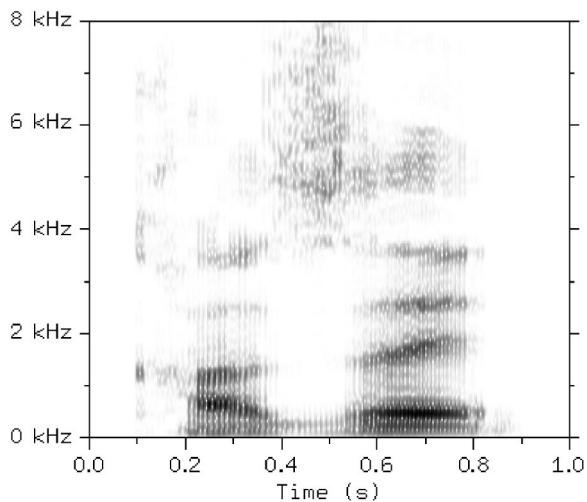


FIG. 12. The spectrogram of the Japanese word “kaze” rate-changed to half-speed.

The M time instants t_i were used as delimiters to segment the scalogram x into a set of $M - 1$ pieces y_i . The segments y_i correspond to individual glottal pulses, similarly sized sections of unvoiced speech, or silence. The notation y_i omits scale levels for simplicity, but all scale levels are implicitly included. If all segments y_i are concatenated in the proper order, $(y_1, y_2, \dots, y_{M-1}) = x$, the original scalogram is recovered. The speech rate was halved by simply duplicating each y_i in the proper sequence, $(y_1, y_1, y_2, y_2, \dots, y_{M-1}, y_{M-1}) = w$. The half-rate waveform z' was then obtained by $S(w) = z'$. A conventional spectrogram of the half-rate waveform is shown in Fig. 12. The speech rate was doubled by concatenating the even-numbered segments $(y_2, y_4, y_6, \dots) = x$, with $S(w) = z'$ as before. The even-numbered segments were chosen because they included the “k” sound; the doubled results from the odd-numbered segments sounded like “aze.” A conventional spectrogram of the double-rate waveform is shown in Fig. 13. Informal listening tests found that both examples were clear and understandable. The realism of the half-rate examples was marred

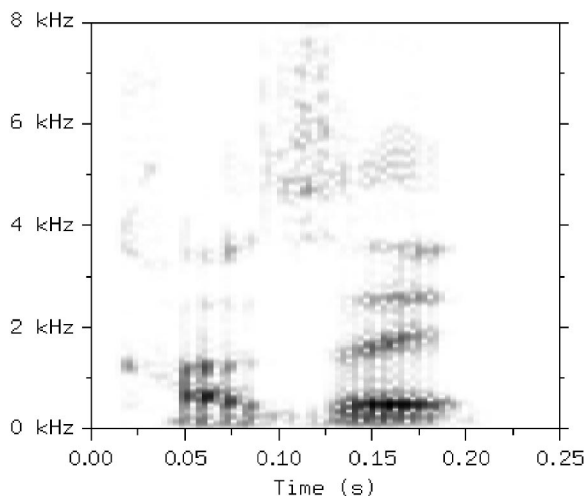


FIG. 13. The spectrogram of the Japanese word “kaze” rate-changed to double speed.

slightly by a mild low-frequency artifact, but the realism of the double-rate example was excellent.

IV. CONCLUSIONS

Procedures for calculating and inverting the SIDWT are described above via equations (Sec. II A) and as a block diagram (Sec. II B). A self-contained proof that the ISIDWT is the pseudoinverse of the SIDWT is given in the Appendix. The new SIDWT/ISIDWT algorithm described here is mathematically equivalent to the stationary wavelet transform and the maximal overlap discrete wavelet transform, and is easier to implement efficiently in an array-oriented mathematical scripting language. The recorded speech example discussed in Sec. III demonstrates that the SIDWT is useful for visualization and analysis of complicated, nonstationary, acoustic waveforms. The SIDWT provides a clear picture of the sounds excited and modulated by the opening and closing of the glottis in speech. The SIDWT is complementary to the STFT for visualization, as they provide optimal views of different aspects of the waveform. The SIDWT and ISIDWT together provide the capability to segment and reconstruct the sound from each individual glottal period for further visualization and analysis. Examples of half-rate and double-rate speech modification demonstrate the potential for the SIDWT/ISIDWT to prove useful for applications that currently use one of the PSOLA family of techniques.

The visualization capabilities of the SIDWT are constrained by the properties of the underlying discrete wavelet transform. In particular, the single-octave bandwidth resulting from dyadic wavelet scaling means that the tonal content of sound waveforms cannot be observed in detail. In certain situations, some tonal information can be extracted by analysis of sum and difference frequencies. The relative lack of frequency-domain information provided by the SIDWT is not really a loss, however; it is a trade-off, which enables the SIDWT to provide more detailed time-domain information. For the same reason, segmentation and reconstruction operations provide more detailed control in the time domain than in the frequency domain. The most productive way to use the SIDWT/ISIDWT will likely prove to be in concert with the STFT and other frequency-domain methods, e.g., the pitch-synchronous spectrogram-like display presented in Sec. III. The SIDWT and ISIDWT are based on purely mathematical principles, and have no inherent connection to psychoacoustics, e.g., two sound waveforms that are “close” in the sense of human perception of sounds are not guaranteed to be close in the sense of the standard Euclidean norm in redundant wavelet coefficient space. The applicability of the SIDWT/ISIDWT to acoustics can only be judged empirically. The experience to date, while subjective, and limited in scope to the field of automotive sound quality and the two rate-changing examples presented here, has been consistently encouraging. Possible speech-related applications include aspects of automated speech recognition and simplified intonation/formant visualization. Editing sound recordings (cutting and pasting sections) with the SIDWT/ISIDWT is especially convenient since the audible clicks and pops produced at section boundaries by simple time-domain editing procedures are automatically prevented. The rate-modified

speech examples presented here show promise, and the SIDWT/ISIDWT may prove useful for other applications that currently employ PSOLA-based methods. Given the broad scope of research in acoustics, there are many potential applications for the SIDWT/ISIDWT in visualization, manipulation, and analysis of nonstationary sound waveforms.

ACKNOWLEDGMENTS

We would like to thank Professor Alan Haddow and Professor William Hartmann of Michigan State University for helpful discussions. We are grateful to Professor Charles MacCluer from the Industrial Mathematics Program at Michigan State University for creating the connections that led to this work.

APPENDIX: PROOF OF THEOREM 1

Our strategy is to prove that T as defined in Sec. II A is an isometry and that S is the adjoint of T . Theorem 1 then follows from a well-known linear algebra fact. According to the definition of T in Sec. II A, we define the mappings

$$T_1: l^2(\mathbb{Z}_N) \ni z \mapsto (x_1, y_1) \in l^2(\mathbb{Z}_{2N}), \quad (\text{A1})$$

and for each stage $j=2, \dots, p$

$$T_j: l^2(\mathbb{Z}_{jN}) \ni (x_1, \dots, x_{j-1}, y_{j-1}) \\ \mapsto (x_1, \dots, x_{j-1}, x_j, y_j) \in l^2(\mathbb{Z}_{(j+1)N}). \quad (\text{A2})$$

Additionally, let $R: l^2(\mathbb{Z}_{(p+1)N}) \mapsto l^2(\mathbb{Z}_{(p+1)N})$ be the reordering given by

$$R(x_1, x_2, \dots, x_p, y_p) = (R_{N/2}(x_1), R_{N/4}(x_2), \dots, \\ R_{N/2^p}(x_p), R_{N/2^p}(y_p)). \quad (\text{A3})$$

Note that the p -stage SIDWT of $z \in l^2(\mathbb{Z}_N)$ is then given by $T(z) = RT_p \cdots T_2 T_1(z)$. Since convolution and reordering are linear maps, each T_j is linear. Therefore, T is a linear transformation.

We will now show that T is an isometry, i.e., that $\langle Tz_1, Tz_2 \rangle = \langle z_1, z_2 \rangle$ for all $z_1, z_2 \in l^2(\mathbb{Z}_N)$. To avoid considering the case $j=1$ separately, in the following let $(x_1, x_0, y_0) = y_0 = z$ and $(x_1, x_0) = 0$.

Theorem 3. The transform T is an isometry.

Proof: We show that for each $j=1, \dots, p$ the mapping T_j is an isometry. Since the reordering operator R is unitary, we have $\langle R w_1, R w_2 \rangle = \langle w_1, w_2 \rangle$ for all $w_1, w_2 \in l^2(\mathbb{Z}_{(p+1)N})$. As a composition of isometries, T is then an isometry.

Let $j \in \{1, \dots, p\}$, $z_1 = (x_1, \dots, x_{j-1}, y_{j-1})$ and $z_2 = (\xi_1, \dots, \xi_{j-1}, \eta_{j-1})$. Then

$$\langle T_j z_1, T_j z_2 \rangle \\ = \left\langle \left(x_1, \dots, x_{j-1}, \frac{1}{\sqrt{2}} R_2(y_{j-1} * \bar{v}), \frac{1}{\sqrt{2}} R_2(y_{j-1} * \bar{u}) \right), \right. \\ \left. \left(\xi_1, \dots, \xi_{j-1}, \frac{1}{\sqrt{2}} R_2(\eta_{j-1} * \bar{v}), \frac{1}{\sqrt{2}} R_2(\eta_{j-1} * \bar{u}) \right) \right\rangle \\ = \langle (x_1, \dots, x_{j-1}), (\xi_1, \dots, \xi_{j-1}) \rangle$$

$$+ \left\langle \frac{1}{\sqrt{2}} R_2(y_{j-1} * \bar{v}), \frac{1}{\sqrt{2}} R_2(\eta_{j-1} * \bar{v}) \right\rangle \\ + \left\langle \frac{1}{\sqrt{2}} R_2(y_{j-1} * \bar{u}), \frac{1}{\sqrt{2}} R_2(\eta_{j-1} * \bar{u}) \right\rangle. \quad (\text{A4})$$

Using the fact that R_2 is unitary, and applying Parseval's relation, we get

$$\left\langle \frac{1}{\sqrt{2}} R_2(y_{j-1} * \bar{v}), \frac{1}{\sqrt{2}} R_2(\eta_{j-1} * \bar{v}) \right\rangle \\ = \frac{1}{2} \langle y_{j-1} * \bar{v}, \eta_{j-1} * \bar{v} \rangle \\ = \frac{1}{2N} \sum_{n=0}^{N-1} \langle \widehat{(y_{j-1} * \bar{v})}(n), \widehat{(\eta_{j-1} * \bar{v})}(n) \rangle \\ = \frac{1}{2N} \sum_{n=0}^{N-1} \langle \hat{y}_{j-1}(n) \hat{v}(n), \hat{\eta}_{j-1}(n) \hat{v}(n) \rangle \\ = \frac{1}{2N} \sum_{n=0}^{N-1} |\hat{v}(n)|^2 \langle \hat{y}_{j-1}(n), \hat{\eta}_{j-1}(n) \rangle. \quad (\text{A5})$$

The same equality holds for \bar{u} instead of \bar{v} . Making use of the known identity $\hat{v}(n) = \hat{v}^*(n)$, we get

$$|\hat{v}(n)|^2 + |\hat{u}(n)|^2 = |\hat{v}(n)|^2 + |\hat{u}(n)|^2 = 2, \quad (\text{A6})$$

since the system matrix $A(n)$ is unitary for all $n=0, \dots, N-1$. This finishes the proof as follows:

$$\langle T_j z_1, T_j z_2 \rangle = \langle (x_1, \dots, x_{j-1}), (\xi_1, \dots, \xi_{j-1}) \rangle \\ + \frac{1}{N} \sum_{n=0}^{N-1} \langle \hat{y}_{j-1}(n), \hat{\eta}_{j-1}(n) \rangle \\ = \langle (x_1, \dots, x_{j-1}), (\xi_1, \dots, \xi_{j-1}) \rangle \\ + \frac{1}{N} \langle \hat{y}_{j-1}, \hat{\eta}_{j-1} \rangle \\ = \langle (x_1, \dots, x_{j-1}), (\xi_1, \dots, \xi_{j-1}) \rangle \\ + \langle y_{j-1}, \eta_{j-1} \rangle \\ = \langle (x_1, \dots, x_{j-1}, y_{j-1}), (\xi_1, \dots, \xi_{j-1}, \eta_{j-1}) \rangle \\ = \langle z_1, z_2 \rangle. \quad \square \quad (\text{A7})$$

Given a linear mapping $L: l^2(\mathbb{Z}_m) \rightarrow l^2(\mathbb{Z}_k)$, $m, k \in \mathbb{N}$, the adjoint operator

$$L^\dagger: l^2(\mathbb{Z}_k) \rightarrow l^2(\mathbb{Z}_m), \quad (\text{A8})$$

is given by the unique mapping defined by the property $\langle Lz, w \rangle = \langle z, L^\dagger w \rangle$ for all $z \in l^2(\mathbb{Z}_m)$, $w \in l^2(\mathbb{Z}_k)$. The matrix corresponding to L^\dagger is just the conjugate transpose of the matrix corresponding to L .

We will show that the ISIDWT S is the adjoint operator of T . To do so, we first prove a lemma.

Lemma. Let $x, y, v \in l^2(\mathbb{Z}_N)$. Then

$$\langle y * \bar{v}, x \rangle = \langle y, x * v \rangle. \quad (\text{A9})$$

Proof: Using Parseval's equality and again the identity $\hat{v}(n) = \hat{v}^*(n)$, we get

$$\begin{aligned} \langle y * \tilde{v}, x \rangle &= \frac{1}{N} \sum_{n=0}^{N-1} \hat{y}(n) \hat{v}(n) \hat{x}^*(n) \\ &= \frac{1}{N} \sum_{n=0}^{N-1} \hat{y}(n) \hat{v}^*(n) \hat{x}^*(n) = \langle y, x * v \rangle, \end{aligned} \quad (\text{A10})$$

which proves that the adjoint of the convolution with v is the convolution with the conjugate reflection \tilde{v} . \square

Following the definition of S in Sec. II A, we define mappings S_j as follows: Let $w = (w_1, w_2, \dots, w_{j+1}) \in l^2(\mathbb{Z}_{(j+1)N})$, where $w_i \in l^2(\mathbb{Z}_N)$ for all $i = 1, \dots, j+1$. Then, for all $j = 1, \dots, p$, we define

$$S_j : l^2(\mathbb{Z}_{(j+1)N}) \ni w \mapsto (w_1, \dots, w_{j-1}, \eta_j) \in l^2(\mathbb{Z}_{jN}), \quad (\text{A11})$$

where

$$\eta_j = \frac{1}{\sqrt{2}} (R_{N/2}(w_{j+1}) * u + R_{N/2}(w_j) * v). \quad (\text{A12})$$

Noting that $S = S_1 S_2 \cdots S_p R^{-1}$, where R^{-1} is the inverse of the reordering operator as defined in Eq. (A3), we can now prove the following theorem.

Theorem 4. $S = T^\dagger$.

Proof: Since $T = R T_p \cdots T_2 T_1$, it follows from the definition of the adjoint that $T^\dagger = T_1^\dagger T_2^\dagger \cdots T_p^\dagger R^\dagger = T_1^\dagger T_2^\dagger \cdots T_p^\dagger R^{-1}$, where we made use of the fact that the adjoint of the unitary operator R is its inverse R^{-1} . Hence, the theorem follows once we show that $T_j^\dagger = S_j$ for all $j = 1, \dots, p$.

We need to prove that $\langle T_j z, w \rangle = \langle z, S_j w \rangle$ for all $z \in l^2(\mathbb{Z}_{jN})$ and $w \in l^2(\mathbb{Z}_{(j+1)N})$. Let $z = (x_1, x_2, \dots, x_{j-1}, y_{j-1})$ and $w = (w_1, w_2, \dots, w_{j+1})$. Then

$$\begin{aligned} \langle T_j z, w \rangle &= \left\langle \left(x_1, x_2, \dots, x_{j-1}, \frac{1}{\sqrt{2}} R_2(y_{j-1} * \tilde{v}), \right. \right. \\ &\quad \left. \left. \frac{1}{\sqrt{2}} R_2(y_{j-1} * \tilde{u}) \right), (w_1, w_2, \dots, w_{j+1}) \right\rangle \\ &= \langle (x_1, x_2, \dots, x_{j-1}), (w_1, w_2, \dots, w_{j-1}) \rangle \\ &\quad + \left\langle \frac{1}{\sqrt{2}} R_2(y_{j-1} * \tilde{v}), w_j \right\rangle \\ &\quad + \left\langle \frac{1}{\sqrt{2}} R_2(y_{j-1} * \tilde{u}), w_{j+1} \right\rangle. \end{aligned} \quad (\text{A13})$$

Note that by the lemma

$$\begin{aligned} \left\langle \frac{1}{\sqrt{2}} R_2(y_{j-1} * \tilde{v}), w_j \right\rangle &= \left\langle y_{j-1} * \tilde{v}, \frac{1}{\sqrt{2}} R_{\frac{N}{2}}(w_j) \right\rangle \\ &= \left\langle y_{j-1}, \frac{1}{\sqrt{2}} R_{\frac{N}{2}}(w_j) * v \right\rangle. \end{aligned} \quad (\text{A14})$$

The same equation holds for u instead of v . So, we conclude that

$$\begin{aligned} \langle T_j z, w \rangle &= \langle (x_1, x_2, \dots, x_{j-1}), (w_1, w_2, \dots, w_{j-1}) \rangle \\ &\quad + \left\langle y_{j-1}, \frac{1}{\sqrt{2}} (R_{\frac{N}{2}}(w_j) * v + R_{\frac{N}{2}}(w_{j+1}) * u) \right\rangle \\ &= \left\langle (x_1, x_2, \dots, x_{j-1}, y_{j-1}), \left(w_1, w_2, \dots, w_{j-1}, \right. \right. \\ &\quad \left. \left. \frac{1}{\sqrt{2}} (R_{\frac{N}{2}}(w_j) * v + R_{\frac{N}{2}}(w_{j+1}) * u) \right) \right\rangle \\ &= \langle z, S_j w \rangle. \square \end{aligned} \quad (\text{A15})$$

We have shown that S is given by T^\dagger . It is well known that a 1–1 linear map L has a unique pseudoinverse given by $(L^\dagger L)^{-1} L^\dagger$. Here, T is not only 1–1 but also an isometry, which is equivalent to $T^\dagger T = id|_{l^2(\mathbb{Z}_N)}$. Thus, the pseudoinverse of T is given by $T^\dagger = S$, which proves Theorem 1.

Instead of u and v in the definition of S , we can use any a and b in $l^2(\mathbb{Z}_N)$, satisfying the condition

$$\hat{a}(n) \hat{a}^*(n) + \hat{b}(n) \hat{b}^*(n) = 2, \quad (\text{A16})$$

for all $n = 0, \dots, N-1$ to obtain a mapping \tilde{S} which is still an inverse of T on the image of T . Since \tilde{S} is computed using convolutions in the same way as S , the number of multiplications required to compute \tilde{S} is the same as for S . However, since then $\tilde{S}^\dagger \neq T$, \tilde{S} is no longer the pseudoinverse, i.e., we lose the notion of ‘‘closeness’’ in the least-squares sense.

¹G. Beylkin, ‘‘On the representation of operators in bases of compactly supported wavelets,’’ SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal. **29**, 1716–1740 (1992).

²R. R. Coifman and D. L. Donoho, ‘‘Translation-invariant de-noising,’’ in *Wavelets and Statistics*, Lecture Notes in Statistics, **103**, edited by A. Antoniadis and G. Oppenheim (Springer, New York, 1995), pp. 125–150.

³H. Sari-Sarraf and D. Brzakovic, ‘‘A shift-invariant discrete wavelet transform,’’ IEEE Trans. Signal Process. **45**, 2621–2626 (1997).

⁴J. Liang and T. W. Parks, ‘‘A translation-invariant wavelet representation algorithm with applications,’’ IEEE Trans. Signal Process. **44**, 225–232 (1996).

⁵D. J. Scholl, ‘‘Translation-invariant data visualization with orthogonal discrete wavelets,’’ IEEE Trans. Signal Process. **46**, 2031–2034 (1998).

⁶G. P. Nason and B. W. Silverman, ‘‘The stationary wavelet transform and some statistical applications,’’ in *Wavelets and Statistics*, Lecture Notes in Statistics, **103**, edited by A. Antoniadis and G. Oppenheim (Springer, New York, 1995), pp. 281–299.

⁷D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis* (Cambridge University Press, Cambridge, UK, 2000), pp. 159–205.

⁸D. Scholl and B. Yang, ‘‘Wavelet-based visualization, separation, and synthesis tools for sound quality of impulsive noises,’’ SAE 2003-01-1527 (The Society of Automotive Engineers, Warrendale, PA, 2003); (<http://www.sae.org>)

⁹C. Duxbury, M. Davies, and M. Sandler, ‘‘Separation of transient information in musical audio using multiresolution analysis techniques,’’ in Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-01) Limerick, Ireland (2001).

¹⁰R. Kronland-Martinet, ‘‘The wavelet transform for analysis, synthesis, and processing of speech and musical sounds,’’ Comput. Music J. **12**, 11–20 (1988).

¹¹D. Darlington, L. Daudet, and M. Sandler, ‘‘Digital audio effects in the wavelet domain,’’ in Proceedings of the 5th Int. Conference on Digital Audio Effects (DAFX-02), Hamburg, Germany (2002).

¹²Y. M. Cheng and D. O’Shaughnessy, ‘‘Automatic and reliable estimation

- of glottal closure instant and period,” *IEEE Trans. Acoust., Speech, Signal Process.* **37**, 1805–1815 (1989).
- ¹³ S. Kadambe and G. F. Boudreaux-Bartels, “Application of the wavelet transform for pitch detection of speech signals,” *IEEE Trans. Inf. Theory* **38**, 917–924 (1992).
- ¹⁴ V. N. Tuan and C. d’Alessandro, “Robust glottal closure detection using the wavelet transform,” in *Proceedings of the European Conference on Speech Technology, EuroSpeech, Budapest, 2805–2808* (1999).
- ¹⁵ I. Hetrick and H. Ackermann, “A vowel synthesizer based on formant sinusoids modulated by fundamental frequency,” *J. Acoust. Soc. Am.* **106**, 2988–2990 (1999).
- ¹⁶ P. J. B. Jackson and C. H. Shadle, “Frication noise modulated by voicing, as revealed by pitch-scaled decomposition,” *J. Acoust. Soc. Am.* **108**, 1421–1434 (2000).
- ¹⁷ A. L. Francis, V. Ciocca, and J. M. C. Yu, “Accuracy and variability of acoustic measures of voicing onset,” *J. Acoust. Soc. Am.* **113**, 1025–1032 (2003).
- ¹⁸ H. M. Hanson and E. S. Chuang, “Glottal characteristics of male speakers: Acoustic correlates and comparison with female data,” *J. Acoust. Soc. Am.* **106**, 1064–1077 (1999).
- ¹⁹ J. C. Lucero and L. L. Koenig, “Time normalization of voice signals using functional data analysis,” *J. Acoust. Soc. Am.* **108**, 1408–1420 (2000).
- ²⁰ M. W. Frazier, *An Introduction to Wavelets through Linear Algebra* (Springer, New York, 1999).
- ²¹ International Phonetic Association (c/o Department of Linguistics, University of Victoria, Victoria, British Columbia, Canada); (<http://web.uvic.ca/ling/resources/ipa/handbook.htm>)
- ²² D. J. Scholl, “Wavelet-based visualization of impulsive and transient sounds in stationary background noise,” SAE 2001-01-1475 (The Society of Automotive Engineers, Warrendale, PA, 2001); (<http://www.sae.org>)
- ²³ P. Mokhtari and K. Tanaka, “A corpus of Japanese vowel formant patterns,” *Bull. of the Electrotechnical Lab. (Electrotechnical Laboratory, Agency of Industrial Science and Technology, Ministry of International Trade and Industry, Japan)*, **64**, 57–66 (2000); (www.etl.go.jp/jp/results/bulletin/pdf/64-rinjji/05tanaka.pdf)
- ²⁴ E. Moulines and J. Laroche, “Nonparametric techniques for pitch-scale and time-scale modification of speech,” *Speech Commun.* **16**, 175–205 (1995).
- ²⁵ M. Holzapfel, R. Hoffmann, and H. Höge, “A wavelet-domain PSOLA approach,” *Third ESCA/COCOSDA Workshop on Speech Synthesis, Blue Mountains, NSW, Australia*, pp. 283–286 (1998).
- ²⁶ R. W. L. Kortekaas and A. Kohlrausch, “Psychoacoustical evaluation of the pitch-synchronous overlap-and-add speech-waveform manipulation technique using single-formant stimuli,” *J. Acoust. Soc. Am.* **101**, 2202–2213 (1997).
- ²⁷ R. W. L. Kortekaas and A. Kohlrausch, “Psychoacoustical evaluation of PSOLA. II. Double-formant stimuli and the role of vocal perturbation,” *J. Acoust. Soc. Am.* **105**, 522–535 (1997).

Defect imaging with guided waves in a pipe

Takahiro Hayashi^{a)} and Morimasa Murase

Faculty of Engineering, Nagoya Institute of Technology, Gokiso Showa Nagoya, 466-8555, Japan

(Received 22 September 2004; revised 6 January 2005; accepted 7 January 2005)

Guided wave techniques are expected to become an effective means for rapid, long-range inspection of pipes. Such techniques still have many practical difficulties in application, however, due to the complex characteristics of guided waves such as dispersion and their multimodal nature. A defect imaging technique is developed in this study to overcome the complexities of guided wave inspection. Received signals are separated into single-mode waveforms with a mode extraction technique and then spatial waveforms on the pipe surface at an arbitrary time are reconstructed. The predicted waveforms can provide a defect image at the moment when an incident wave arrives at a defect region, which is based on a time-reversal technique. This defect imaging technique is experimentally verified using eight signals detected at eight different circumferential positions. Images of artificial defects are obtained with one-hole and two-hole test pipes, and increasing the frequency of incident waves increases the resolution of the images. Holes and pipe ends are recognizable in the images, but the reconstructed images contain some errors in the area behind the defects where guided waves do not propagate or do not reflect back to the receiving transducers. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1862572]

PACS numbers: 43.60.Lq, 43.60.Fg, 43.20.Mv [YHB]

Pages: 2134–2140

I. INTRODUCTION

A nondestructive evaluation (NDE) technique for pipe-work has been developed using guided waves acting as ultrasonic modes propagating longitudinally in a pipe. Guided wave NDE is expected to become a rapid long-range inspection technique^{1,2} in which an inspection tool emits guided waves from an accessible part of a pipe, then detects reflected or transmitted guided waves from damaged areas of the pipe that are inaccessible (such as buried or coated regions). Because guided waves propagate as forward or backward traveling waves, transducers lined up on the circumference of a pipe can fully detect reflected waves or transmitted waves. Reflected and transmitted waves contain all of the information about reflective walls and scattering objects. If the detected signals are processed properly, most of this information can be acquired. This characteristic of guided waves is highly beneficial for NDE when considering that not all energy due to refraction and scattering can be detected by NDE using bulk waves.

Guided waves in a pipe, however, have multimodal and dispersive characteristics similar to Lamb waves.^{3,4} These distinctive characteristics make it difficult to evaluate pipe damage and defects with sufficient accuracy and reliability using conventional pulse-echo and pitch-catch techniques to measure amplitude and velocity. For example, the phase and amplitude of signals received from a transducer will often be completely different even at one wavelength away in the longitudinal direction. This is because guided waves are a superposition of many modes having different wavelengths and wave structures. And, because these characteristics of guided waves depend on diameter, thickness, and the material properties of the pipe, the results obtained from one pipe test cannot be directly applied to other types of pipes.

Easy-to-use axisymmetric modes have therefore been adopted for guided wave inspections. Axisymmetric modes have a wide nondispersive region in the dispersion curves, particularly an axisymmetric torsional mode $T(0,1)$ having constant phase and group velocities over all frequency regions. As such, incident pulse signals of axisymmetric modes maintain their waveforms and propagate long distances with a high signal-to-noise ratio. Also, single axisymmetric modes can be detected easily using axisymmetric transducers with no mixing of nonaxisymmetric modes.

However, because axisymmetric modes propagate no information concerning the circumferential direction (such as circumferential location, maldistribution, or size of defects) and defect characterization cannot be accomplished in an axisymmetric mode, several studies on nonaxisymmetric modes of guided waves in a pipe have been recently presented. Ditri⁵ investigated scattering waves from a circumferential crack using the S -parameter formalism and addressed the possibility of crack sizing. Lowe *et al.*⁶ and Alleyne *et al.*⁷ analyzed mode converted guided waves from part-circumferential notches in a pipe both experimentally, using the mode extraction technique, and theoretically, using a finite element model. Li and Rose⁸ experimentally verified the angular profile of guided waves that were estimated by normal mode expansion theory, and determined the position of transducers for defect detection. Demma *et al.*⁹ investigated the reflected guided waves from a part-circumferential notch when a torsional axisymmetric mode is emitted, both experimentally and through FE analysis.

In this paper we describe a defect imaging technique, in which nonaxisymmetric modes, as well as an axisymmetric mode, are considered and an image of defects is obtained after appropriate signal processing. Guided wave propagation at an arbitrary time and point on the surface of a pipe is predicted by a combination of the mode extraction technique

^{a)}Corresponding author; electronic mail: hayashi@nitech.ac.jp

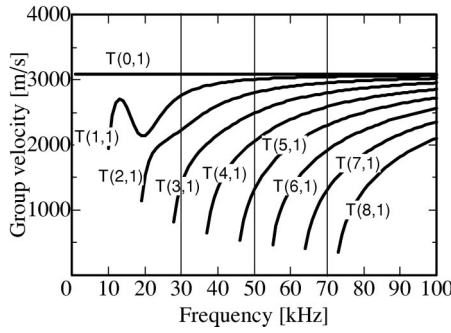


FIG. 1. Group velocity dispersion curves of torsional modes for an aluminum pipe with outer diameter=111 mm and thickness=3.5 mm.

with the reconstruction of spatial waveforms. This technique is based on the time-reversed method proposed by Ing and Fink³ in the sense that received waveforms are played back, but mode extraction and the reconstruction process are required.

The mode extraction technique, which is based on Lowe's method but includes negative circumferential orders, is described first, and defect imaging is described in detail. This imaging technique is then experimentally verified for one-hole and two-hole pipes at various frequencies.

II. THEORY

A. Circumferential mode extraction

The normal mode expansion technique gives the displacement on the surface of a pipe for a harmonic wave $\exp(-i\omega t)$ using the cylindrical coordinate system (r, θ, z) as

$$u(\theta, z, t) = \sum_{n=-\infty}^{+\infty} \sum_{m=1}^{+\infty} \alpha_{nm} \exp(in\theta + ik_{nm}z - i\omega t), \quad (1)$$

where n and α_{nm} represent the circumferential order number (family) and the amplitude of the m th mode in the n th family, respectively.¹⁰⁻¹² Wave numbers k_{nm} consist of several propagating modes with real values and an infinite number of nonpropagating (evanescent) modes having complex values.¹²

Torsional modes $T(n, m)$ dominated mainly by a displacement in the circumferential direction are now considered. For additional details about indices n and m , the reader should refer to Ref. 13, and to the dispersion curves of Fig. 1. Normally, guided wave NDE uses propagating modes with real wave numbers k_{nm} . For axisymmetric torsional modes $T(0, m)$, the minimum frequencies above which the m th propagating modes occur (the cutoff frequencies) are given for general-purpose pipes when the pipe wall thickness is small compared to the radius as

$$f_c = \frac{c_T}{2d}(m-1), \quad (2)$$

where c_T and d are the transverse wave velocity and the thickness of the pipe, respectively. General-purpose pipes, such as schedule pipes, are at most approximately 15 mm thick with a transverse wave velocity of approximately 3200 m/s. The cutoff frequency for $m=2$ is therefore greater than 106 kHz in Eq. (2). When we consider the frequency range

from 20 to 100 kHz, used widely in long-range pipe inspections, modes greater than $m=2$ can be neglected. The displacement on the surface of a pipe in Eq. (1) can be rewritten as

$$u(\theta, z, t) = \sum_{n=-\infty}^{+\infty} \alpha_n \exp(in\theta + ik_n z - i\omega t), \quad (3)$$

where $\alpha_{n1} \equiv \alpha_n$ and $k_{n1} \equiv k_n$, in order to simplify the representation.

Signals received in a small section $r_0 d\theta$ at $z=z_R$ and $\theta=\theta$ are given by $u(\theta, z_R, t)r_0 d\theta$. Multiplying a weight function $\exp(-in_E\theta)$ against the term $u(\theta, z_R, t)r_0 d\theta$, and integrating this with respect to θ , gives extracted waveforms of the n_E th circumferential mode as

$$\begin{aligned} u_{n_E}^{\text{ext}}(z_R, t) &= \int_0^{2\pi} u(\theta, z_R, t) \exp(-in_E\theta) r_0 d\theta \\ &= 2\pi r_0 \alpha_{n_E} \exp(ik_{n_E} z_R - i\omega t). \end{aligned} \quad (4)$$

In practice, receiving sensors are not infinitesimal as indicated by $r_0 d\theta$, and the number of receiving points is finite. Assuming N receiving positions in the circumferential direction at regular intervals as

$$\theta_k = \frac{2\pi}{N}(k-1), \quad (5)$$

and assuming that a displacement is detected in the region with an aperture of $\theta_0 = 2\pi/N$, then the received signals $\theta = \theta_k$ and $z = z_R$ are given by

$$\begin{aligned} u^R(\theta_k, z_R, t) &= \int_{\theta_k - \theta_0/2}^{\theta_k + \theta_0/2} u(\theta, z_R, t) r_0 d\theta \\ &= r_0 \sum_{n=-\infty}^{+\infty} \alpha_n f_n(\theta_0) \exp(in\theta_k + ik_n z_R - i\omega t), \end{aligned} \quad (6)$$

where

$$f_n(\theta_0) = \begin{cases} \theta_0, & \text{for } n=0, \\ \frac{2 \sin(n\theta_0/2)}{n}, & \text{for } n \neq 0. \end{cases} \quad (7)$$

Then, multiplying by the weight function $\exp(-in_E\theta_k)$ and summing with respect to k gives, similar to Eq. (4), the extracted waveforms

$$\begin{aligned} u_{n_E}^{\text{ext}}(z_R, t) &= r_0 \sum_{k=1}^N \sum_{n=-\infty}^{+\infty} \alpha_n f_n(\theta_0) \\ &\quad \times \exp\{i(n-n_E)\theta_k + ik_n z_R - i\omega t\} \\ &\approx r_0 \alpha_{n_E} f_{n_E}(\theta_0) \exp(ik_{n_E} z_R - i\omega t). \end{aligned} \quad (8)$$

It should be noted here that Eq. (8) holds under the condition of $|n_E| \leq N/2 - 1$, because modes lower than n_E are superposed in Eq. (8) for $|n_E| \geq N/2$. In dispersion curves for torsional modes, those modes higher than $n=8$ do not propagate in the frequency regions used in this study (30–70 kHz). Also, for the highly dispersive modes of $n=4-7$, the detected signals become very broad, and only a very small part

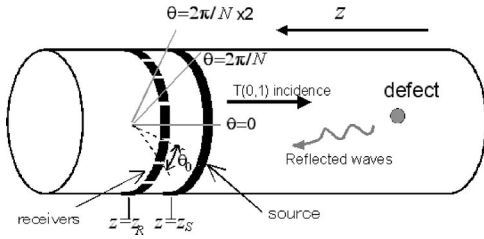


FIG. 2. Schematic setup for defect imaging.

of the energy of these modes can be recorded. Therefore, eight different circumferential positions are used to detect the guided waves in later experiments. This corresponds to $N = 8$, and the main propagating modes from $n = -3$ to $+3$ can be extracted.

B. Waveform prediction of dispersive modes and reconstruction of defect images

If the diameter, thickness, and material properties of a pipe are given, then the wave numbers k_{n_E} and phase velocities c_{n_E} of the extracted modes can be theoretically obtained. Once the wave numbers k_{n_E} are given, multiplying $\exp\{ik_{n_E}(z-z_R)\}$ against the extracted waveforms $u^{\text{ext}}(z_R, t)$ obtained from received signals at $z=z_R$ yields predicted waveforms of the n_E th mode at $z=z$ as

$$u_{n_E}^{\text{ext}}(z_R, t) \exp\{ik_{n_E}(z-z_R)\} \approx r_0 \alpha_{n_E} f_{n_E}(\theta_0) \exp(ik_{n_E}z - i\omega t) = u_{n_E}^{\text{ext}}(z, t). \quad (9)$$

Equation (9) is for a harmonic wave of frequency ω . In actual experiments, extracted waveforms in the frequency domain are obtained by FFT first, and then a series of data multiplied by $\exp\{ik_{n_E}(z-z_R)\}$ are transformed back to the time domain by inverse FFT. This process was presented as a dispersion reduction technique for dispersive Lamb waves by Wilcox *et al.*¹⁴

From Eq. (6), waveforms received with uniform weight in the region having an aperture of θ_0 at an arbitrary point (θ, z) on the surface of a pipe are given by

$$u^R(\theta, z, t) = r_0 \sum_{n=-\infty}^{+\infty} \alpha_n f_n(\theta_0) \exp(in\theta + ik_n z - i\omega t). \quad (10)$$

Using the predicted extracted waveforms $u_n^{\text{ext}}(z, t)$ obtained by Eq. (9), Eq. (10) can be rewritten as

$$u^R(\theta, z, t) = \sum_{n=-\infty}^{n=+\infty} u_n^{\text{ext}}(z, t) \exp(in\theta). \quad (11)$$

This equation shows that multiplying $\exp(in\theta)$ against $u_n^{\text{ext}}(z, t)$ obtained in Eq. (9) gives waveforms detected by a receiving transducer with aperture θ_0 at an arbitrary point.

The fact that reflected waves from defects are detected at N different circumferential positions (Fig. 2) is now considered. A $T(0,1)$ mode is emitted at $z=z_S$ toward the defects in the $-z$ direction. The $T(0,1)$ incident wave is reflected back from the defects in the $+z$ direction, and the reflected waveforms are detected by N receiving transducers located at z

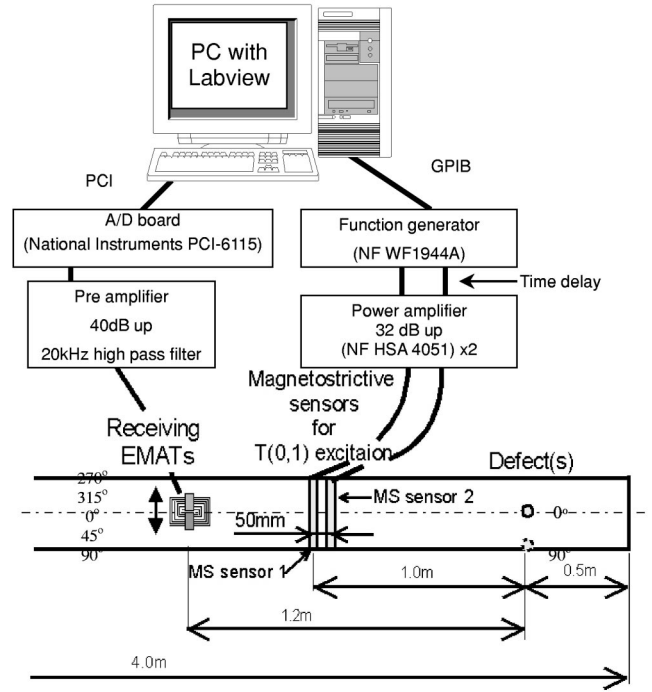


FIG. 3. Experimental setup.

$=z_R$. From the N received signals, spatial waveforms at an arbitrary time can be predicted like snapshots at all of the time steps. Since the reflected waveforms are very large at the moment of reflection from a reflective object, the snapshot at this moment shows an image of the reflective object.

III. EXPERIMENT

A. Experimental setup

Figure 3 shows a schematic figure of the experimental setup used in this study. A magnetostrictive sensor (MS sensor) consisting of a nickel plate adhered to the pipe surface and a coil wound on the nickel plate, is used for generating a $T(0,1)$ mode.¹⁵ The nickel plates are magnetized in the circumferential direction and serve as bias magnetic fields. Alternating voltage as four-cycle sinusoidal waves of 30, 50, and 70 kHz are applied to the coil, generating vibration in the circumferential direction by magnetostriction. Two magnetostrictive sensors are used to double the $T(0,1)$ mode signals by applying a time delay to one forward-located sensor (MS sensor2 in Fig. 3). A function generator (NF corporation, WF1944A) and a power amplifier (NF corporation, HSA4051) are used to apply rf signals to the coils.

An electromagnetic acoustic transducer (EMAT) is adopted for noncontact measurements. The experiments of this study require accurate waveform measurements at many different circumferential positions; therefore, noncontact measurement with EMATs is more suitable than contact methods strongly affected by contact pressure and the thickness of the coupling material. The EMAT used in the experiment consists of two permanent magnets (NEOMAX, NEOMAX Co. Ltd., 15 mm \times 6 mm \times 10 mm) and sheet coils fabricated by a printed circuit technique. The EMAT is flexible so that it fits the curvature of the pipe. The sheet coil and

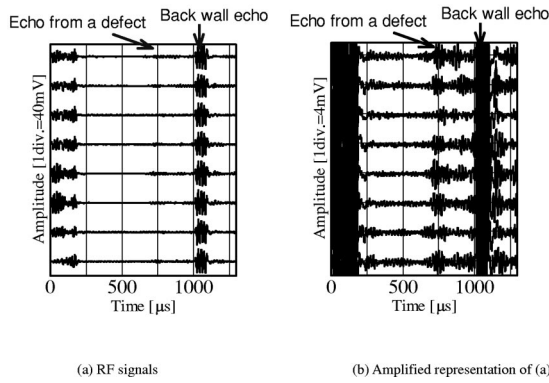


FIG. 4. Waveforms detected at eight different circumferential positions. (a) rf signals; (b) amplified representation of (a).

permanent magnets of the EMAT are arranged to receive shear horizontal waves by Lorentz forces. This transducer, then, receives vibration mainly in the circumferential direction.¹⁶ In order to receive displacement in the region of 1/8 of the circumference, the EMAT is designed to be 44 mm in width. Signals received at the EMAT are magnified 40 dB by a preamplifier and are recorded after 10 signal averagings.

The test objects for this study are aluminum pipes of 4 m in length, 111 mm in diameter, and 3.5 mm in thickness with a longitudinal wave velocity $c_L=6260$ m/s and a transverse wave velocity $c_T=3080$ m/s. The defects are located 0.5 m from the pipe end and one meter from MS sensor1 (Fig. 3). Backward waves may propagate in a direction opposite to the defects, but in this study the test pipes are large enough to separate the necessary signals from the unwanted signals that come from the other side of the pipe.

The receiving EMAT is positioned 1.2 m from the defects and measures signals at eight different circumferential positions. The artificial defects are through-holes of 10 mm in diameter and are located at 0° for the one-hole test pipe and at 0° and 90° for the two-hole test pipe, in the circumferential direction, as shown in Fig. 3.

Figures 4(a) and 4(b) are the waveforms detected at eight different circumferential positions (0° , 45° , 90° , 135° , 180° , 225° , 270° , and 315°) for a one-hole test pipe and 50 kHz incidence. Large signals around $1000 \mu\text{s}$ are backwall echoes from the pipe end 1.5 and 1.7 m in distance from excitation MS sensor1 and the receiving EMAT, respectively. Small signals between 500 and $1000 \mu\text{s}$ are echoes from the defect. Figure 4(b) shows the waveforms of Fig. 4(a) after amplification to discern the waveforms of the defect echoes. The backwall echoes and incident waves around $t=0 \mu\text{s}$ are unnecessary for defect imaging, so all signals from 330 to $980 \mu\text{s}$ are extracted, and other amplitudes are assumed to be zero.

B. Prediction of waveforms reflected at a pipe defect

Figure 5 shows spatial waveforms at three different times predicted by the above process using the detected waveforms of Fig. 4. All images show absolute values of spatial waveforms normalized by the maximum amplitude of spatial waveforms at all time steps. Since (a) $t=195 \mu\text{s}$ is a time when the incident T(0,1) mode has not yet arrived at the

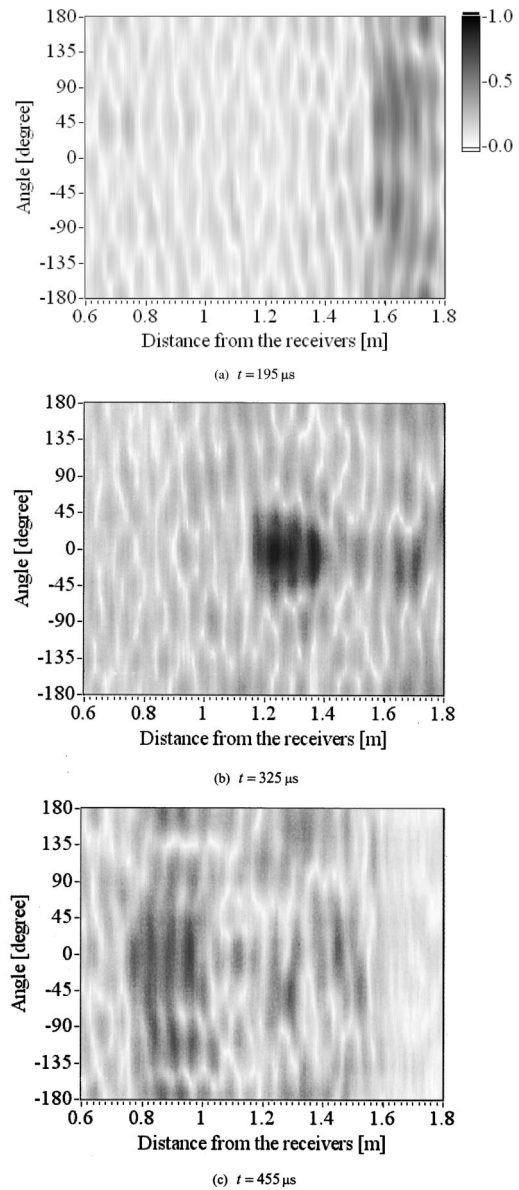


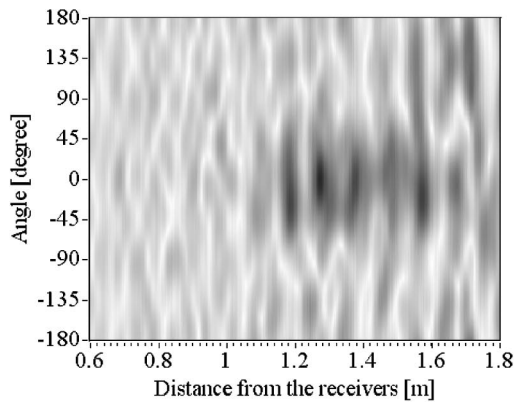
FIG. 5. Snapshots of predicted waveforms at various moments. (a) $t=195 \mu\text{s}$, (b) $t=325 \mu\text{s}$, (c) $t=455 \mu\text{s}$.

defect, the predicted waveform (a) does not give a defect image. In (b), the moment when the incident T(0,1) mode arrives at the defect, a distinct spot with high intensity is present at approximately 1.2 m from the receivers and 0° , which agrees well with the position of the artificial defect. After (b), the waveforms spread out from the defect, as can be seen in (c).

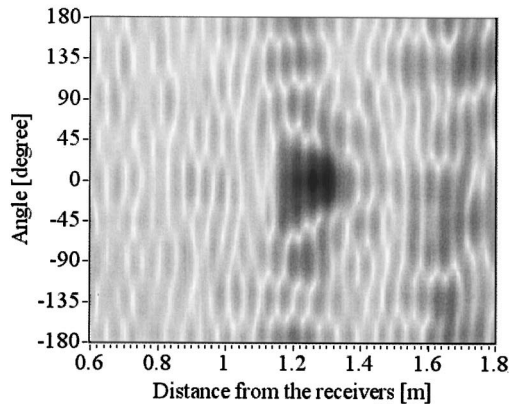
C. Effect of frequency on the predicted waveforms

Next, snapshots of the predicted waveforms in the one-hole pipe at $t=325 \mu\text{s}$ are shown in Fig. 6 for a four-cycle sinusoidal incidence of 30 and 70 kHz. For (a) 30 kHz, the high-intensity spot becomes broad in both the longitudinal and circumferential directions, indicating lower resolution. On the other hand, a better result with higher resolution is obtained in both directions for (b) 70 kHz.

Since resolution in the longitudinal direction is determined by a wave disturbance zone, then in the case of four-



(a) 30 kHz, $t = 325 \mu\text{s}$



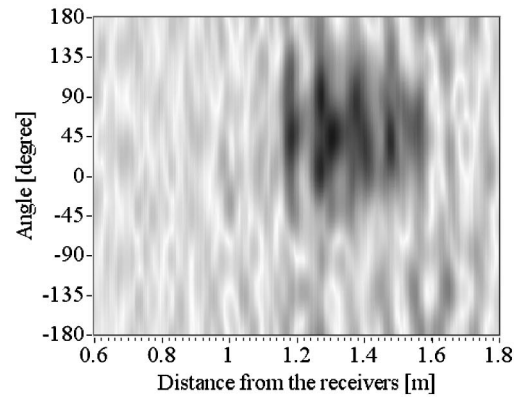
(b) 70 kHz, $t = 325 \mu\text{s}$

FIG. 6. Predicted waveforms for various frequencies at $t = 325 \mu\text{s}$. (a) 30 kHz, $t = 325 \mu\text{s}$; (b) 70 kHz, $t = 325 \mu\text{s}$.

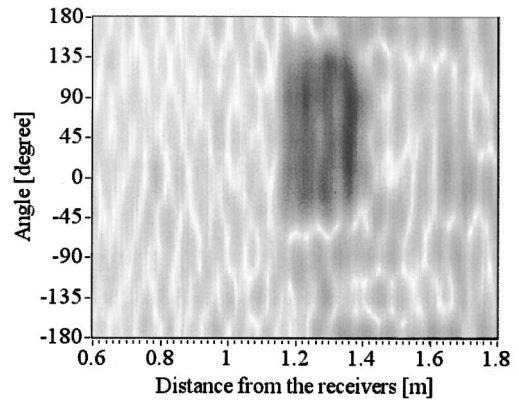
cycle incidence, four times the wavelength is taken as the high-intensity zone. Circumferential resolution is determined by the order of detectable circumferential modes. In the frequency region of approximately 30 kHz, modes of orders higher than T(3,1) cannot be detected due to the cutoff frequency, and the T(2,1) and T(3,1) modes have large dispersion. It is expected, therefore, that signals within the given gate ($t = 330$ to $980 \mu\text{s}$) do not contain all of the energy of the T(2,1) and T(3,1) modes, causing low resolution in the circumferential direction. In contrast, in the high-frequency region of approximately 70 kHz, since higher-order modes such as T(2,1) and T(3,1) have low-dispersion characteristics, most of the energy of the higher-order modes can be obtained, resulting in higher resolution in the circumferential direction.

Figure 7 shows snapshots at the moment when the incident T(0,1) mode arrives at the defects ($t = 325 \mu\text{s}$) for a two-hole test pipe with through-holes at 0° and 90° in the three different frequency regions. In the 30 and 50 kHz defect images, single broad dark regions can be seen, the centers of which are located at about 45° . In the 70 kHz defect image, however, two distinct dark spots can be seen at 0° and 90° , at 1.2 m.

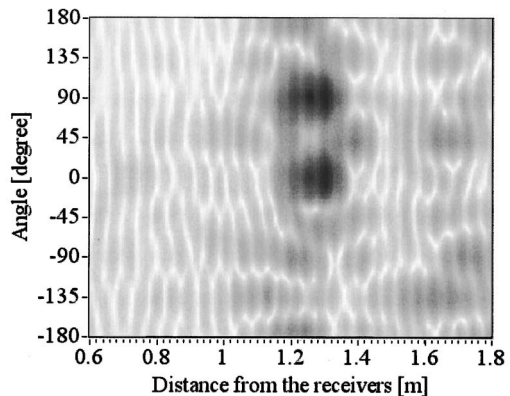
The dark zones even at 70 kHz for both one-hole and two-hole test pipes are shown to be larger than the artificial holes of 10 mm in diameter. We can expect that the dark



(a) 30 kHz



(b) 50 kHz



(c) 70 kHz

FIG. 7. Predicted waveforms at $t = 325 \mu\text{s}$ for a two-hole test pipe and various frequencies. (a) 30 kHz, (b) 50 kHz, (c) 70 kHz.

zones become closer to the circle 10 mm in diameter at higher frequencies. However, since the predicted waveform, as shown in Eq. (10), denotes the waveform at (θ, z, t) measured by the receiving transducer with an aperture of $\theta_0 = 2\pi/N$, the resolution in the circumferential direction is theoretically limited to the aperture of receiving transducers $\theta_0 = 2\pi/N$.

D. Defect imaging considering the arrival region of the incident wave

In Figs. 5–7, which show the waveforms predicted from the received signals, the defect images are obtained as snapshots at the moment that the incident T(0,1) mode arrives at

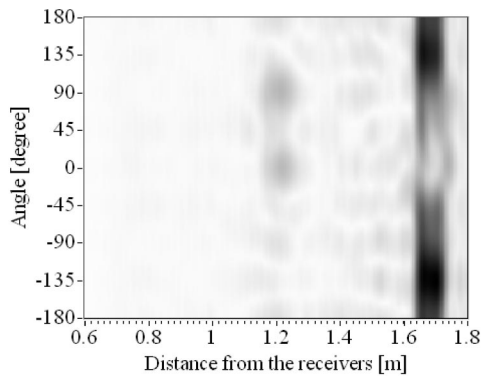


FIG. 8. Imaging of defects and a pipe end.

a defect region. However, when reflection objects are located at different longitudinal locations, information on the reflection objects in the longitudinal direction cannot be obtained in one snapshot. In this section we describe defect imaging in which such reflection objects are shown in one image. Summing up the images of predicted waveforms at all time steps only in the arrival region of incident wave $T(0,1)$ provides an image of reflection objects, such as defects, pipe ends, weld lines, and branches. Acquiring an image of the two holes and pipe end using the received signals that produced the snapshot image of Fig. 7(c) is now considered. The image in Fig. 7(c) was obtained by gating the received signals between 330 and 980 μs in order to avoid backwall echoes. In the following process, however, the signals are gated between 330 and 1200 μs in order to include reflection waves from the pipe end.

Now, consider the arrival region of an incident $T(0,1)$ mode at a certain time t , which is from $z = c_T t - l\lambda$ to $z = c_T t$, where l and λ are the number of cycles and the wavelength of the incident burst waves, respectively. If reflective walls, such as defects, exist in this region, then the predicted waveforms of the reflected waves become large. In contrast, if there are no reflective objects present, then the predicted waveforms become zero. Thus, summing up the absolute values of the predicted waveforms in the regions for all time steps gives an image of the reflective objects. This is shown in Fig. 8, in which a reflective object with high intensity is present at approximately 1.7 m from the receivers, and two round reflective objects with low intensity are present at 1.2 m, $\theta=0^\circ$ and $\theta=90^\circ$. The two dark regions at 1.2 m agree well with the two artificial holes. The dark region at 1.7 m should indicate the pipe ends, but is not symmetric with respect to the pipe axis and has a nonreflecting region behind the two holes. It is considered that this region cannot be reconstructed as an image due to scattering at the holes and multiple reflections between the holes and the pipe end. However, it is confirmed that this imaging technique has a potential for providing images of reflective objects having longitudinal distributions.

E. Possible errors by measurement conditions

The imaging technique using guided waves presented herein is applicable when guided wave modes can be perfectly separated and dispersion curves are theoretically ob-

tained. However, it is possible to superpose unwanted modes and to change a defect image due to the characteristic of the receivers. In this study, we assume that the EMATs detect only horizontal shear vibration by Lorentz force, but the EMATs can also detect unexpected longitudinal or out-of-plane vibrations. In contrast, if we can separately detect vibrations in all directions and extract flexural modes as well as torsional modes, better defect images can be obtained.

Random noise, largely detected as shown in Fig. 4, may also distort defect images. However, random noise interfering with received signals is reduced in the summation with respect to k in Eq. (8). Therefore, the dark defect zone in Fig. 5(b) appears large compared with other noisy regions. This indicates the robustness of the defect imaging technique against random noise interfering with each waveform.

In this technique, however, the sensitivity of the transducers at all receiving points should coincide because all signals are added and subtracted in the imaging process. Since we use one EMAT for receiving at eight positions in this study, this error must be small, but when using the contact technique, in which it is difficult to control contact pressure and the thickness of the coupling medium, sensitivity mismatching could result in large distortion in defect images. In order to investigate the influence of differences in sensor sensitivity, Fig. 9 shows the defect images when specific signals cannot be properly obtained, for a two-hole test pipe and 70 kHz incidence. Signals were gated from 330 to 980 μs in order to obtain only images of defects and to avoid pipe ends. The defect image in Fig. 9(a) was constructed by using all of the signals, the same signals that were used to obtain Figs. 7(c) and 8. Figures 9(b) and 9(c) were obtained by replacing the first signals received at $\theta=0^\circ$ and the first two signals at $\theta=0^\circ$, 45° with zero-valued samples, respectively, showing the typical case when a portion of the signals cannot be detected due to transducer trouble in practical testing. In Fig. 9(d), only the first signals at $\theta=0^\circ$ are phase-reversed and the other signals are the same, indicating an insufficient contact condition for the contact-type transducers, or a difference in sensor sensitivity. Comparing Figs. 9(a), 9(b), and 9(c), as the number of zero signals increases, a more unwanted dark zone can be seen in the images. In Fig. 9(d), a very large distortion of the defect images can be seen, although only one waveform was reversed. These results reveal that differences in transducer sensitivity, as well as a lack of signals, may cause large image distortion.

IV. CONCLUSIONS

Defect imaging with guided waves was described. A defect image in a pipe was obtained as a spatial waveform by processing signals detected at many different circumferential positions. After a detailed theoretical description, the defect imaging technique was experimentally verified using eight signals detected at eight different circumferential points.

In defect imaging tests for 30, 50, and 70 kHz four-cycle sinusoidal waves, higher resolutions in both the longitudinal and circumferential directions were obtained at higher frequencies. Reconstructed images for two-hole and one-hole pipes agreed well with the test pipes. Moreover, a pipe end

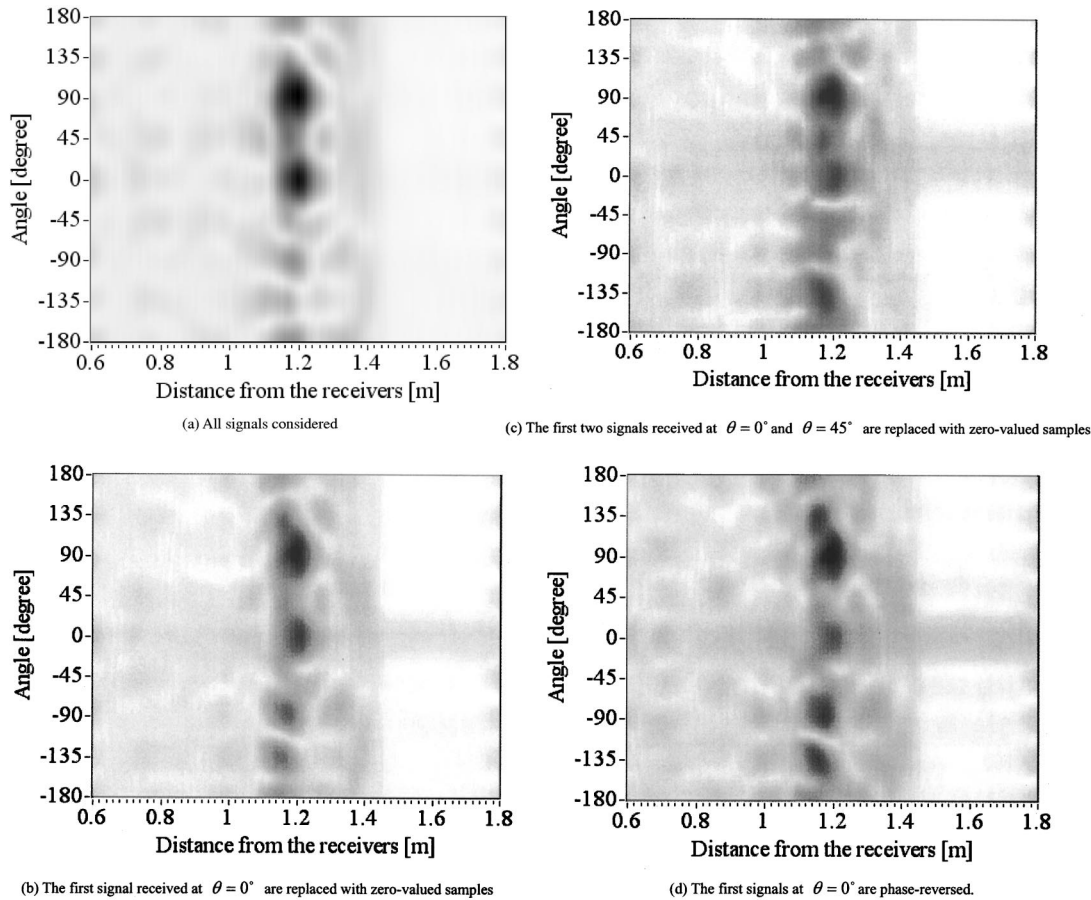


FIG. 9. Defect images when specific signals cannot be properly obtained, for a two-hole test pipe and 70 kHz incidence. (a) All signals considered. (b) The first signals received at $\theta=0^\circ$ are replaced with zero-valued samples. (c) The first two signals received at $\theta=0^\circ$ and $\theta=45^\circ$ are replaced with zero-valued samples. (d) The first signals at $\theta=0^\circ$ are phase-reversed.

behind two holes could be seen in the image, but an undetectable region was found just behind the holes.

In this study, we used only torsional modes and shear horizontal vibration. From a theoretical point of view, the imaging technique described herein can be applied to longitudinal and flexural modes. Because flexural modes in reflected waves from defects were detectable, combining this technique with other modes should enhance the defect images.

ACKNOWLEDGMENTS

This research was supported by the Industrial Technology Research Grant Program in 2003 from the New Energy and Industrial Technology Development Organization (NEDO) of Japan. We thank Takashi Ikeda at the CXR Corporation for his considerable assistance with the magnetostrictive sensors.

- ¹J. L. Rose, "A baseline and vision of ultrasonic guided wave inspection potential," *J. Pressure Vessel Technol.* **124**, 273–282 (2002).
- ²P. Cawley, M. J. S. Lowe, D. N. Alleyne, B. Pavlakovic, and P. Wilcox, "Practical long range guided wave inspection—applications to pipes and rail," *Mater. Eval.* **61**, 67–74 (2003).
- ³R. K. Ing and M. Fink, "Time reversed Lamb waves," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**, 1032–1043 (1998).
- ⁴T. Hayashi and J. L. Rose, "Guided wave simulation and visualization by a semi-analytical finite element method," *Mater. Eval.* **61**, 75–79 (2003).
- ⁵J. J. Ditri, "Utilization of guided elastic waves for the characterization of

- circumferential crack in hollow cylinders," *J. Acoust. Soc. Am.* **96**, 3769–3775 (1994).
- ⁶M. J. S. Lowe, D. N. Alleyne, and P. Cawley, "The mode conversion of a guided wave by a part-circumferential notch in a pipe," *J. Appl. Mech.* **65**, 649–656 (1998).
- ⁷D. N. Alleyne, M. J. S. Lowe, and P. Cawley, "The reflection of guided waves from circumferential notches in pipes," *J. Appl. Mech.* **65**, 635–641 (1998).
- ⁸J. Li and J. L. Rose, "Excitation and propagation of non-axisymmetric guided waves in a hollow cylinder," *J. Acoust. Soc. Am.* **109**, 457–468 (2001).
- ⁹A. Demma, P. Cawley, M. Lowe, and A. G. Roosenbrand, "The reflection of the fundamental torsional mode from cracks and notches in pipes," *J. Acoust. Soc. Am.* **114**, 611–625 (2003).
- ¹⁰D. C. Gazis, "Three-dimensional investigation of the propagation of waves in hollow circular cylinders. I. Analytical foundation. II. Numerical Results," *J. Acoust. Soc. Am.* **31**, 568–578 (1959).
- ¹¹J. L. Rose, *Ultrasonic Waves in Solid Media* (Cambridge University Press, Cambridge, 1999).
- ¹²T. Hayashi, K. Kawashima, Z. Sun, and J. Rose, "Analysis of flexural mode focusing by a semi-analytical finite element method," *J. Acoust. Soc. Am.* **113**, 1241–1248 (2003).
- ¹³H. Nishino, S. Takashina, F. Uchida, M. Takemoto, and K. Ono, "Modal analysis of hollow cylindrical guided waves and applications," *Jpn. J. Appl. Phys.* **40**, 364–370 (2001).
- ¹⁴P. D. Wilcox, M. J. S. Lowe, and P. Cawley, "A signal processing technique to remove the effect of dispersion from guided wave signals," *Rev. Prog. Quant. Nondestr. Eval.* **20**, 555–562 (2001).
- ¹⁵H. Kwun, S. Y. Kim, and G. M. Light, "The magnetostrictive sensor technology for long range guided wave testing and monitoring of structures," *Mater. Eval.* **61**, 80–84 (2003).
- ¹⁶M. Hirao and H. Ogi, *EMATs For Science and Industry—Noncontacting Ultrasonic Measurements* (Kluwer, Dordrecht, 2003).

Simplified nonlinear outer hair cell models

Niranjan Deo^{a)}

Department of Mechanical Engineering, 2350 Hayward Avenue, University of Michigan, Ann Arbor, Michigan 48109-2125

Karl Grosh^{b)}

Department of Mechanical and Department Biomedical Engineering, 2350 Hayward Avenue, University of Michigan, Ann Arbor, Michigan 48109-2125

(Received 29 September 2004; revised 24 January 2005; accepted 25 January 2005)

We present a consistent second-order expansion of nonlinear constitutive theories for outer hair cells. For a particular theory, we will test the validity of such a model for small variations in voltage and strain about the resting state of outer hair cells. An analysis of the various terms in the simplified nonlinear model and their relevance to outer hair cell mechanics are presented. Results show that the second-order expansion is adequate for modeling outer hair cell mechanics in a global model of the cochlea. Model predictions agree with the notion that voltage nonlinearities are the dominant ones at low sound levels *in vivo*. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1871753]

PACS numbers: 43.64.Bt, 43.64.Kc [BLM]

Pages: 2141–2146

I. INTRODUCTION

Outer hair cells (OHCs) along with their stereocilia are believed to be the primary source of nonlinearity in the mammalian cochlea, contributing to its sensitivity and frequency selectivity. These electromotile cells play an important role in generating electrically evoked emissions, distortion product emissions, and influence, to a significant degree, other experimental artifacts seen in the cochlea. For modeling the cochlea it is hence essential to capture OHC behavior. There is a wealth of data in the literature on *in vitro* experiments on OHCs. Results from Parthasarathi *et al.* (2003) indicate that OHC sensitivity to electrical and mechanical resting conditions seen *in vitro* are also important *in vivo*. If we assume that the *in vitro* behavior of OHCs is a fair representation of *in vivo* mechanics, we can proceed to extract information from *in vitro* data to construct models for *in vivo* mechanics.

While there have been attempts in the literature to construct complete OHC models (for instance, Tolomeo and Steele, 1995; Raphael *et al.*, 2000; Spector, 2001; Iwasa, 2001; Deo and Grosh, 2004), none of the current models explain all features of OHC electromotility seen in experiments. For cochlear models the *in vivo* behavior of OHCs is of primary interest. OHCs are subjected to a much narrower range of voltages and strains *in vivo* than in *in vitro* experiments. In this paper, we show how to construct and analyze an OHC model containing first-order nonlinearities. Such a model, while unable to predict OHC response over the range of voltages used in *in vitro* experiments, possesses sufficient richness to describe *in vivo* nonlinearities and response. Furthermore, a first-order model will be easier to incorporate in a global cochlear model and also easier to construct from experiments. The extended area motor model (EAMM) (Deo and Grosh, 2004), which is an extension upon Iwasa's original hypothesis (Iwasa, 1994; Iwasa, 2001), is used as a basis

to construct a simpler model. The EAMM incorporates the most recent experimental findings of voltage-dependent OHC stiffness (He and Dallos, 1999; He and Dallos, 2000) as well as the nonlinear capacitance and motility. The simple model is derived from an energy basis to ensure reciprocity. The dynamic response of OHCs seen in experiments is successfully reproduced by the simple model. The model parameters are related to experiments to illustrate their relevance to OHC behavior.

II. ENERGY-BASED OHC MODELS

OHCs are shell-like hollow cylindrical bodies filled with fluid. The OHC is assumed to be a right-circular cylinder of length L , radius r , and membrane thickness h (see Fig. 1). The applied loads are considered axisymmetric and it is assumed that the principal axis of the material properties coincides with the cylindrical coordinates. All bending stresses on the OHC are neglected. Since the radius to thickness ratio of the OHC is very high, all radial strains are ignored. The motor proteins are assumed to be spread uniformly on the cylindrical OHC surface. These assumptions leave us with charge (Q), voltage (V), axial strain (ϵ_z), axial stress (σ_{zz}), circumferential strain (ϵ_c), and circumferential stress ($\sigma_{\theta\theta}$) as the state variables. Instead of stresses we use tensions, $T_z = \sigma_{zz}h$ and $T_c = \sigma_{\theta\theta}h$, for the forcing state variables to have similarity with the area motor model. Consider the Gibbs free energy per unit surface area of the OHC membrane:

$$G = G(T_z, T_c, V). \quad (1)$$

Experimentally, it is easier to measure strain rather than tension so an energy function with voltage and strains as the independent variables will be more useful. Using the Legendre transformation we obtain

$$W(\epsilon_z, \epsilon_c, V) = G(T_z, T_c, V) + T_z \epsilon_z + T_c \epsilon_c. \quad (2)$$

The corresponding charge and tensions are

^{a)}Electronic mail: ndeo@umich.edu

^{b)}Electronic mail: grosh@umich.edu

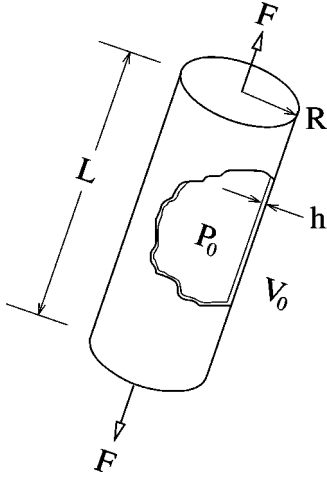


FIG. 1. Schematic of an outer hair cell having length L , radius r , and thickness h . F is the external force acting on the OHC. P_0 is the initial fluid pressure and V_0 is the transmembrane voltage.

$$Q = - \frac{\partial W(\varepsilon_z, \varepsilon_c, V)}{\partial V}, \quad (3)$$

$$T_z = \frac{\partial W(\varepsilon_z, \varepsilon_c, V)}{\partial \varepsilon_z}, \quad (4)$$

$$T_c = \frac{\partial W(\varepsilon_z, \varepsilon_c, V)}{\partial \varepsilon_c}. \quad (5)$$

We next specialize to the most common *in vitro* experimental configuration, an intact cylindrical OHC with forcing along its cylindrical axis. For a cylindrical OHC with a fluid pressure P_0 inside the cell, the force equilibrium relations can be written as

$$T_z = \frac{RP_0}{2} + T_{\text{ext}}, \quad (6)$$

$$T_c = RP_0, \quad (7)$$

where T_{ext} is the external axial tension acting on the OHC [$T_{\text{ext}} = F/(2\pi R)$]. After the OHC is pressurized, it is assumed that the cell undergoes incompressible deformation, imposing a constraint on subsequent deformations $\varepsilon_z + 2\varepsilon_c = \varepsilon_{\text{vol}}$; where ε_{vol} , a constant, is the strain due to initial pressurization of the cell (see Iwasa, 1993; Iwasa and Adachi 1997; and Deo and Grosh, 2004; for more details). Since ε_z and ε_c are related, we can replace ε_c in our expressions for energy, stress, and charge. For the rest of the paper, the axial strain ε_z is referred to as simply ε and external tension as T instead of T_{ext} . After the substitution we can write the charge and external load as

$$\frac{\partial W(\varepsilon, V)}{\partial V} = -Q, \quad (8)$$

$$\frac{\partial W(\varepsilon, V)}{\partial \varepsilon} = T_z - \frac{T_c}{2} = T. \quad (9)$$

III. CONSISTENT SECOND-ORDER EXPANSION

The free-energy functions, $W(\varepsilon, V)$, typically have a complex form, dependent on exponentials (Spector, 2000) or

Boltzmann functions (Deo and Grosh, 2004; Iwasa, 2001). We are interested in extracting a simpler OHC model from such energy functions, which works near the resting state. Let V_0 , Q_0 , T_0 , and ε_0 denote voltage, charge, tension, and strain, respectively, at the resting state. Let $q = Q - Q_0$, $t = T - T_0$, $v = V - V_0$, and $\varepsilon = \varepsilon - \varepsilon_0$ denote the shift from the values at a resting state. Expanding Eqs. (8) and (9) in Taylor series about their resting state yields

$$q = - \frac{\partial^2 W}{\partial V^2} \Big|_{\varepsilon_0, V_0} v - \frac{\partial^2 W}{\partial V \partial \varepsilon} \Big|_{\varepsilon_0, V_0} \varepsilon - \frac{1}{2} \frac{\partial^3 W}{\partial V^3} \Big|_{\varepsilon_0, V_0} v^2 - \frac{\partial^3 W}{\partial V^2 \partial \varepsilon} \Big|_{\varepsilon_0, V_0} v \varepsilon - \frac{1}{2} \frac{\partial^3 W}{\partial V \partial \varepsilon^2} \Big|_{\varepsilon_0, V_0} \varepsilon^2 + \mathcal{O}(\varepsilon^3, \varepsilon^2 v, \varepsilon v^2, v^3), \quad (10)$$

$$t = \frac{\partial^2 W}{\partial \varepsilon^2} \Big|_{\varepsilon_0, V_0} \varepsilon + \frac{\partial^2 W}{\partial \varepsilon \partial V} \Big|_{\varepsilon_0, V_0} v + \frac{1}{2} \frac{\partial^3 W}{\partial \varepsilon^3} \Big|_{\varepsilon_0, V_0} \varepsilon^2 + \frac{\partial^3 W}{\partial \varepsilon^2 \partial V} \Big|_{\varepsilon_0, V_0} \varepsilon v + \frac{1}{2} \frac{\partial^3 W}{\partial \varepsilon \partial V^2} \Big|_{\varepsilon_0, V_0} v^2 + \mathcal{O}(\varepsilon^3, \varepsilon^2 v, \varepsilon v^2, v^3). \quad (11)$$

All second-order terms are retained and higher-order terms are neglected. The question to be answered is whether these are sufficient or do we need the higher-order terms. The above expressions can be written in terms of more familiar variables,

$$q = Cv + e\varepsilon + \frac{1}{2} C_2 v^2 + C_1 v \varepsilon - \frac{1}{2} k_1 \varepsilon^2, \quad (12)$$

$$t = k\varepsilon - ev + \frac{1}{2} k_2 \varepsilon^2 + k_1 \varepsilon v - \frac{1}{2} C_1 v^2. \quad (13)$$

In the above expressions [Eqs. (12), (13)] the constants multiplying linear terms are $C = -\partial^2 W / \partial V^2 \Big|_{\varepsilon_0, V_0}$, which is a capacitance-like term, $k = \partial^2 W / \partial \varepsilon^2 \Big|_{\varepsilon_0, V_0}$ a stiffness-like term, and $e = -\partial^2 W / \partial V \partial \varepsilon \Big|_{\varepsilon_0, V_0} = -\partial^2 W / \partial \varepsilon \partial V \Big|_{\varepsilon_0, V_0}$ the electro-mechanical coupling term. The constants multiplying the nonlinear terms are $k_1 = \partial^3 W / \partial V \partial \varepsilon^2 \Big|_{\varepsilon_0, V_0} = \partial^3 W / \partial \varepsilon^2 \partial V \Big|_{\varepsilon_0, V_0}$, $C_1 = -\partial^3 W / \partial V^2 \partial \varepsilon \Big|_{\varepsilon_0, V_0} = -\partial^3 W / \partial \varepsilon \partial V^2 \Big|_{\varepsilon_0, V_0}$, $k_2 = \partial^3 W / \partial \varepsilon^3 \Big|_{\varepsilon_0, V_0}$, and $C_2 = -\partial^3 W / \partial V^3 \Big|_{\varepsilon_0, V_0}$. These quantities can be directly derived from the energy function W or from experiments if the experiments are done with the appropriate parameters held constant. For example, the “stiffness” k can be obtained by keeping the voltage constant and measuring change in tension T as strain ε is changed. The slope of the curve at a certain strain will give “ k ” at that strain and voltage. These expansions will have a limited range of validity as the higher-order terms start dominating at large strains and voltages away from resting values. Note, the free-energy function corresponding to the second-order expansions has cubic terms which can give rise to negative energy if strains and voltage deviate too greatly from equilibrium.

IV. SIMPLIFIED EAMM

For the extended area motor model (Deo and Grosh, 2004), the Gibbs free energy of an OHC is

$$G = -\frac{1}{2}s_z^l T_z^2 - \frac{1}{2}s_c^l T_c^2 - T_z T_c s_{zc}^l - \frac{1}{2}C_{\text{lin}} V^2 - n T_z a_z - n T_c a_c - n q V + n k_B \theta \log(P_e) + G_0, \quad (14)$$

where s_z^l , s_c^l , and s_{zc}^l are compliance coefficients of the motor in long state, a_z and a_c are anisotropic area changes for the motor occurring during state transition in the axial and circumferential directions, respectively, n is the motor density, q is the charge transferred per motor during change of state, C_{lin} is the linear capacitance of the motor, P_e is the Boltzmann probability function governing the change of states, k_B is Boltzmann's constant, θ is the temperature, and G_0 is a constant (see Deo and Grosh, 2004). Specifying the transmembrane voltage and internal turgor pressure at the resting state determines the volume constraint for the EAMM. Once the operating conditions for the OHC (i.e., the external force and transmembrane voltage) have been determined, a simplified version of the extended area motor model can be derived using Eqs. (2), (10), and (11). Any perturbations in strain and voltage will then take place about these resting values as the transmembrane voltage and external force change.

Analytical expressions for the various constants ($C, k, e, C_1, k_1, C_2, k_2$) were obtained from the energy expression for the EAMM for simulating the simple model. Since the expressions are rather long, they have not been reproduced here. The following first-order nonlinear model is for an OHC having a resting transmembrane voltage of -55 mV and resting strain of -0.47% . The OHC is $70 \mu\text{m}$ long and $10 \mu\text{m}$ in diameter. The resting state corresponds to an equal distribution of motors in the long and short state (i.e., $P_e = 0.5$). There is no external force, and turgor pressure at the resting state is 200 Pa. The expressions for the perturbed force and charge are

$$q = 2.117 \times 10^{-2} v - 3.145 \times 10^{-3} \epsilon + 2.713 \times 10^{-2} v^2 + 3.652 \times 10^{-2} v \epsilon + 2.087 \times 10^{-2} \epsilon^2, \quad (15)$$

$$t = 7.201 \times 10^{-3} \epsilon + 3.145 \times 10^{-3} v + 2.247 \times 10^{-2} \epsilon^2 - 4.173 \times 10^{-2} \epsilon v - 1.826 \times 10^{-2} v^2. \quad (16)$$

Here q is in Coulombs, v is in Volts, and t is in N m^{-1} . The constants and their units are also listed in Appendix A.

V. RESULTS

A. Stiffness and capacitance

The total OHC stiffness and capacitance are given by

$$K_{\text{OHC}} = \frac{2\pi R}{L} (k + k_1 v + k_2 \epsilon), \quad (17)$$

$$C_{\text{OHC}} = 2\pi R L (C + C_1 \epsilon + C_2 v). \quad (18)$$

The simplified model yields stiffness and capacitance that are linearly dependent on the strain and voltage. Figure 2 shows stiffness and capacitance versus voltage for three different

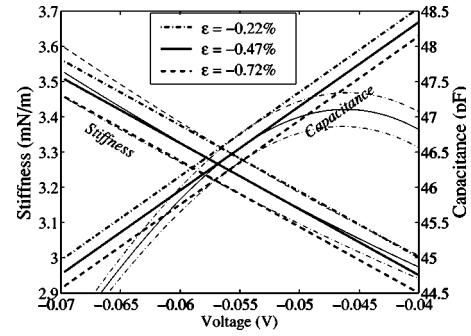


FIG. 2. Stiffness and capacitance versus voltage for three different strains (thick lines). Parameters for the simplified model [Eqs. (15) and (16)] have been selected to match the Deo and Grosh (2004) model at resting state $V = -55$ mV and strain of -0.47% . Results from the Deo and Grosh (2004) model are superimposed as thin lines of the same line type. Note the close match in stiffness for the selected range of voltage and strain.

strains. The results from the full EAMM (thin lines) are also superimposed for comparison.

The simplified model captures the stiffness nonlinearity with respect to voltage with a relative error of less than 1.5%. The capacitance is not captured as well since the operating point is near the peak of the capacitance–voltage relation. However, the maximum relative error in capacitance ($\approx 3\%$) is similar to the relative error in stiffness over a 15 mV range of voltage away from the resting voltage of -55 mV. Figure 3 shows stiffness and capacitance as strain is varied for three different voltages.

For a small range of voltage and strain the simple model does capture the overall features of the OHC behavior. In the next section we compare dynamic simulations of the model with experiments to check whether the first-order nonlinear terms are sufficient to generate the nonlinearity seen in experiments.

B. Dynamic simulations

Dynamic experiments on OHCs (He and Dallos, 1999; He and Dallos, 2000) indicate strong nonlinearity with respect to voltage and weak (if any) nonlinearity with respect to strain. In the experiments the OHCs were excited using a glass probe with known stiffness at one frequency (f_1) and a

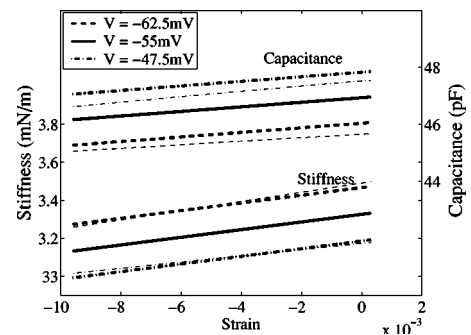


FIG. 3. Stiffness and capacitance versus strain for three different voltages (thick lines). Parameters for the simplified model [Eqs. (15) and (16)] have been selected to match Deo and Grosh (2004) model at resting state $V = -55$ mV and strain of -0.47% . Results from the Deo and Grosh (2004) model are superimposed as thin lines of the same line type. Note the close match in stiffness for the selected range of voltage and strain.

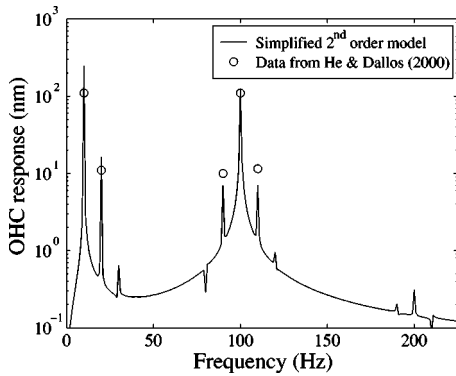


FIG. 4. Frequency spectrum of OHC response. Simultaneous excitation by mechanical and electrical input. The voltage stimulus is at 10 Hz with 40 mV amplitude while the probe signal is at 100 Hz with 400 nm amplitude. Circles represent data from Fig. 3F in He and Dallos (2000). Note the match at sum and difference frequencies (90 and 100 Hz). Probe stiffness used in the model is 4.73 mN/m.

transmembrane voltage signal at another frequency (f_2). For a voltage signal amplitude of 40 mV at 10 Hz and a probe displacement of amplitude 400 nm at 100 Hz the simple model shows dynamic characteristics similar to what are seen in experiments for same conditions (see Fig. 4). The solution procedure is described in Appendix B.

Figure 4 indicates that the second-order expansion possesses sufficient richness to capture experimentally seen behavior. Contributions arising from higher-order terms will be much smaller than contributions from the second-order terms included in this model and can be ignored.

C. Analysis of terms in the model

1. Further simplifications

The simple model includes all possible voltage and strain nonlinearities up to the second order and they contribute equally to the OHC response for voltage and strain signals of similar amplitude. In the dynamic experiments the amplitude of the voltage signal “ v ” was approximately seven times the strain signal given through the probe (0.04 V versus strain of 5.7×10^{-3} for a $70 \mu\text{m}$ long OHC). Therefore both the dynamic simulation and the experiments do not show any significant harmonics of the mechanical signal. Depending on relative amplitudes of voltage and strain signals, a further simplification of the model may be possible by neglecting nonlinear terms corresponding to the weaker signal. Consider the expression for a change in the external tension “ t ,”

$$t = k\epsilon - ev + \frac{1}{2}k_2\epsilon^2 + k_1\epsilon v - \frac{1}{2}C_1v^2. \quad (19)$$

The term that will produce the sum and difference tones is $k_1\epsilon v$; term $k_2\epsilon^2$ will generate harmonics of the mechanical signal while the remaining nonlinear term C_1v^2 will generate harmonics of the voltage signal. Figure 5 shows effect of neglecting these nonlinear terms on the dynamic response of the model for the same excitation as in Fig. 4. Since we have very weak mechanical harmonics for this experimental setup, ignoring the $k_2\epsilon^2$ term does not affect the overall response. Such simplifications of the first-order nonlinear model will

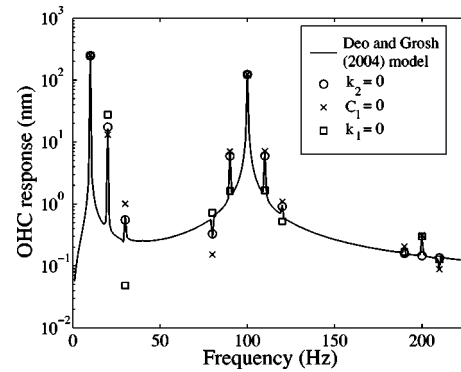


FIG. 5. Frequency spectrum of OHC response for different model settings. The figure shows the effect of neglecting the nonlinear terms in the model. The terms have been removed one at a time, i.e., when $k_1=0$; other constants have their original value.

depend on the range of strains and voltages the cell will be subjected to *in vivo*.

2. Data on strain and voltage *in vivo*

To model *in vivo* OHC activity amplitudes of voltage and strain *in vivo* are needed. We have significant amount of data on strain versus voltage from *in vitro* experiments; for instance Santos-Sacchi (1991, Fig. 2) indicates a change in strain about 0.7% for a 30 mV voltage change in the absence of external forcing. However, *in vitro* data on electromotile strain should not be relied on to get estimates for *in vivo* strain directly, as we do not have any estimates of the forces acting on the OHCs at a resting state *in vivo*. The voltage-dependent stiffness implies that if there is any OHC preloading *in vivo*, a change in the receptor potential will cause the OHC to undergo strain due to the stiffness change. This strain will be in addition to the electromotile strain caused by the voltage change. The effect of stiffness change is discussed in detail by Parthasarathi *et al.* (2003).

Amplitudes of *in vivo* strain can be estimated from vibration measurements of the organ of Corti. Khanna and Hao (1999) report vibration amplitude of around 400 nm for the reticular lamina (RL) at the third row outer hair cell location for a sound level input of around 90 dB in a temporal bone preparation. The amplitude drops linearly to 40 nm for a 70 dB sound pressure level (SPL). The data was for motion at the apical portion of a guinea-pig cochlea. Vibration levels of the basilar membrane (BM) at the apical end have been found to be several hundred times lower than the RL and in phase with RL motion up to the characteristic frequency of the region (Khanna and Hao, 2000). So if we assume the RL motion amplitude to be a good representation of third row OHC amplitude, that would give us a maximum strain of around 0.0057 (for a $70 \mu\text{m}$ OHC) at 90 dB SPL and maximum strain of around 0.00057 at 70 dB. Estimates of the *in vivo* receptor potential are as high as 35 mV for high-intensity sounds (Dallos, 1985) although a more typical value is around 10 mV (see Fig. 6 in Dallos, 1985). The alternating receptor potential saturates at high sound pressure levels as seen in the small difference in receptor potential at 70 dB SPL and 90 dB SPL (Fig. 6 in Dallos, 1985). If we make a conservative estimate that the receptor potential at

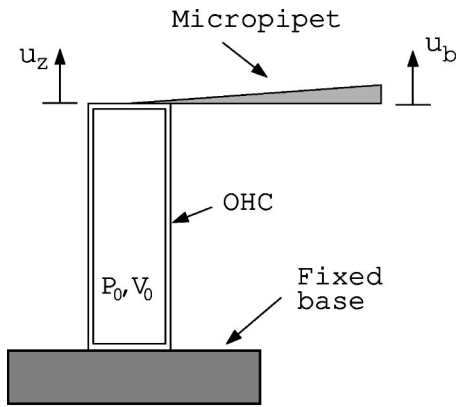


FIG. 6. OHC experiment: initially pressurized, stimulated by applied voltage about the resting potential. Mechanical loading is through a micropipet (modeled as a linear spring). Resting internal fluid pressure P_0 and transmembrane voltage V_0 .

high sound levels is 10 mV, then this would give us a voltage to strain amplitude ratio of 1.75 for 90 dB and 17.5 for 70 dB. Further complicating matters, it is the transmembrane voltage (receptor potential minus extracellular potential) that drives the OHC motility. Experimental measurements of these quantities, which are notoriously difficult (Kossl and Russell, 1992) also yield similar, but lower estimates. This data seems to indicate that at low sound levels, voltage non-linearity will overwhelm any mechanical nonlinearity. If this is true then we can make further simplifications by ignoring the second-order term in ϵ in Eq. (13). Note, however, that the ϵ^2 term has to be retained in the expression for q [Eq. (12)] to maintain reciprocity.

3. Effect of external loading on force-driven motility

The constants in the model are evaluated at a given state of transmembrane voltage, external force, and turgor pressure. It is of interest to investigate how these will vary as the conditions change based on the EAMM. The predictions depend on the model chosen. Experiments can then be used to validate if the model predictions are correct. For instance, if an external compressive force is applied on the OHC, the EAMM predicts the stiffness (k) to decrease and the slope (k_1) of the voltage-dependent stiffness to increase. The k_1 term is the one responsible for the sum-difference tones (see Fig. 5). While the stiffness (k) is always positive, the k_1 term crosses over from negative to positive at a certain external load. The resting turgor pressure in the cell determines the external load at which the slope (k_1) of the voltage-dependent stiffness changes sign. An OHC under compression would therefore exhibit lower stiffness and higher force-driven motility while the amplitude of sum-difference tones will depend on the resting turgor pressure in the cell. Experiments by He and Dallos show that the amplitude of the sum and difference tones can increase or decrease with an increase in cell compression (see Fig. 11 in He and Dallos, 2000). The EAMM predicts the absolute value of k_1 (the sum-difference tone generating term) to decrease and then increase with an increase in compression and therefore is consistent with the two plots in He and Dallos (2000). Since we do not know the exact turgor pressure in the cells, more

of such experimental data showing variation of modulation components in response to increasing compression will help confirm if the EAMM shows the correct behavior. According to the EAMM we should not observe a rise in modulation components followed by a fall in response to increased compression. He and Dallos (2000) also claim to observe both an increase and decrease in force-driven motility with increasing compression. This fact is difficult to reconcile with our current understanding of how stiffness of the OHC changes. For a fixed transmembrane voltage, any external compression should drive the cell into either hyperpolarization or depolarization and therefore to have higher stiffness or lower stiffness, respectively. Variability of turgor pressure cannot explain this phenomenon, as the turgor pressure is just another form of loading. Unless some other factors like “ q ,” the charge transferring across the cell wall also depend on the forcing, this issue is as of yet unresolved.

VI. DISCUSSION

The precise way in which the different parts of the organ of Corti interact is still not clear. Are the OHCs pretensioned at the resting state? Data from Mammano and Ashmore (1993) shows large contraction of OHCs in an isolated cochlear turn preparation when current was passed across the partition. Mammano and Ashmore (1993) state that the OHCs are depolarized by the current and therefore they contract. This will be true if the OHCs are under compression at the resting state, since depolarized OHCs will contract and will have lower stiffness causing them to yield more against compressive forces. On the other hand, if the OHCs are under tension then hyperpolarized OHCs will also exhibit contraction provided shortening due to the increased stiffness dominates any lengthening due to hyperpolarization. This point has also been elaborated upon in Parthasarathi *et al.* (2003). Information on how the tectorial membrane (TM) moves is also very important. Khanna *et al.* (1998) show TM motion to be in phase with RL motion for an apical location (low frequency) and TM velocity to be slightly higher than RL velocity. However, to gauge when stereocilia channels are opening and when they are closing we also need information on the relative RL-TM displacement direction. These questions need to be answered to elaborate on the role played by OHCs in amplifying basilar membrane motion.

The experimental data available on OHCs in literature are of limited use for determining the constants in the simple model introduced here. Consider the capacitance experiments for instance. These are usually performed with no external load on the OHC. So if we construct a model with external force and voltage as the independent variables, then such experiments help yield the capacitance and derivative of the capacitance with respect to voltage at a constant voltage and zero external force. To determine the remaining constants the variation of strain with respect to voltage at a constant external force (zero), and the variation of charge and strain with respect to force at a constant voltage are needed. Except for Dong *et al.* (2002), which looked for reciprocity in a linearized OHC model, such experiments have not been attempted.

TABLE I. Parameter values used for the model.

Parameter	Value	Unit
C	2.117×10^{-2}	$\text{C m}^{-2} \text{V}^{-1}$
C_1	3.652×10^{-2}	$\text{C m}^{-2} \text{V}^{-1}$
C_2	5.426×10^{-2}	$\text{C m}^{-2} \text{V}^{-2}$
e	-3.145×10^{-3}	C m^{-2} or $\text{N m}^{-1} \text{V}^{-1}$
k	7.201×10^{-3}	N m^{-1}
k_1	-4.173×10^{-2}	$\text{N m}^{-1} \text{V}^{-1}$
k_2	4.493×10^{-2}	N m^{-1}

VII. CONCLUSION

The simplified model discussed in this paper is based on a more elaborate model given by Deo and Grosh (2004). There are other models in the literature for OHC electromotility besides the area motor model, for instance, the membrane bending model (Raphael *et al.*, 2000), which assumes flexoelectricity as the basis for OHC electromotility. The second-order expansion presented here will work equally well with such energy-based models.

In a global cochlear model, the OHC representation needs to be accurate only over a small range of voltage and strain seen *in vivo*. As such, these simplified models can be used to represent OHC behavior in global cochlear models. The various parameters of the model can be determined directly from experiments on OHCs or from an approximation of more detailed models, as was done in this paper.

The simplified second-order model provides a good match to OHC behavior for small changes in voltage and strain, provided operating voltages are away from the maxima in the capacitance to voltage relation (see Fig. 2). If it is desired to capture higher fluctuations, a fully nonlinear theory may be needed. A simpler alternative would be to include a third-order voltage term in the charge equation [Eq. (12)], which will give a parabolic dependence of the capacitance on voltage (not shown here). Simulations with such a term included in the model demonstrate an excellent match with the EAMM.

If strains are an order of magnitude smaller than voltage fluctuations, then even simpler expressions can be obtained by ignoring second-order terms in strain. The resulting model will have only voltage-dependent stiffness as is predicted by *in vitro* experiments (He and Dallos, 2000).

APPENDIX:

A. Parameter values

Parameters obtained from the extended area motor model (Deo and Grosh, 2004). Values are for no external forcing, a resting potential of -55 mV, and a resting turgor pressure of 200 Pa (corresponding axial strain is 0.47%). (See Table I.)

B. Dynamic simulations

The experimental setup for the dynamic simulations is shown in Fig. 6. The micropipet spring is in series with the outer hair cell (OHC) spring. Therefore the force acting on the OHC is equal to the stiffness of the micropipet multiplied by its deflection. The displacement of the top of the micropi-

pet spring and the OHC transmembrane voltage are specified. Then using Eq. (13) we can solve for the strain ϵ . Figures 4 and 5 show the Fast Fourier Transform (FFT) of the strain response.

ACKNOWLEDGMENT

This work was supported by National Institutes of Health Grant No. NIDCD R01-04084.

- Dallos, P. (1985). "Response characteristics of mammalian cochlear hair-cells," *J. Neurosci.* **5**, 1591–1608.
- Deo, N., and Grosh, K. (2004). "Two state model for outer hair cell stiffness and motility," *Biophys. J.* **86**, 3519–3528.
- Dong, X., Ospeck, M., and Iwasa, K. H. (2002). "Piezoelectric reciprocal relationship of the membrane motor in the cochlear outer hair cell," *Biophys. J.* **82**, 1254–1259.
- He, D. Z. Z., and Dallos, P. (1999). "Somatic stiffness of cochlear outer hair cells is voltage-dependent," *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8223–8228.
- He, D. Z. Z., and Dallos, P. (2000). "Properties of voltage-dependent somatic stiffness of cochlear outer hair cells," *J. Assoc. Res. Otolaryngol.* **01**, 64–81.
- Iwasa, K. H. (1993). "Effect of stress on the membrane capacitance of the auditory outer hair cell," *Biophys. J.* **65**, 492–498.
- Iwasa, K. H. (1994). "A membrane motor model for the fast motility of the outer hair cell," *J. Acoust. Soc. Am.* **96**, 2216–2224.
- Iwasa, K. H. (2001). "A two-state piezoelectric model for outer hair cell motility," *Biophys. J.* **81**, 2495–2506.
- Iwasa, K. H., and Adachi, M. (1997). "Force generation in the outer hair cell of the cochlea," *Biophys. J.* **73**, 546–555.
- Khanna, S., and Hao, L. (2000). "Amplification in the apical turn of the cochlea with negative feedback," *Hear. Res.* **149**, 55–76.
- Khanna, S., Ulfendahl, M., and Steele, C. (1998). "Vibration of reflective beads placed on the basilar membrane," *Hear. Res.* **116**, 71–85.
- Khanna, S. M., and Hao, L. F. (1999). "Reticular lamina vibrations in the apical turn of a living guinea pig cochlea," *Hear. Res.* **132**, 15–33.
- Kossl, M., and Russell, I. (1992). "The phase and magnitude of hair cell-receptor potentials and frequency tuning in the guinea-pig cochlea," *J. Neurosci.* **12**, 1575–1586.
- Mammano, F., and Ashmore, J. F. (1993). "Reverse transduction measured in the isolated cochlea by laser michelson interferometry," *Nature (London)* **365**, 838–841.
- Parthasarathi, A. A., Grosh, K., Zheng, J., and Nuttall, A. L. (2003). "Effect of current stimulus on *in vivo* cochlear mechanics," *J. Acoust. Soc. Am.* **113**, 442–452.
- Raphael, R. M., Popel, A. S., and Brownell, W. E. (2000). "A membrane bending model of outer hair cell electromotility," *Biophys. J.* **78**, 2844–2862.
- Santos-Sacchi, J. (1991). "Reversible inhibition of voltage-dependent outer hair cell motility and capacitance," *J. Neurosci.* **11**, 3096–3110.
- Spector, A. A. (2000). "Thermodynamic potentials and constitutive relations for a nonlinear electroelastic biological membrane," in *Symposium on the Mechanics and Electromagnetic Materials and Structures*, edited by J. Yang and G. A. Maugin (IOS Press, Amsterdam).
- Spector, A. A. (2001). "A nonlinear electroelastic model of the auditory outer hair cell," *Int. J. Solids Struct.* **38**, 2115–2129.
- Tolomeo, J. A., and Steele, C. R. (1995). "Orthotropic piezoelectric properties of the cochlear outer hair cell wall," *J. Acoust. Soc. Am.* **97**, 3006–3011.

Modeling high-frequency electromotility of cochlear outer hair cell in microchamber experiment

Zhijie Liao^{a)} and Aleksander S. Popel

Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland 21205

William E. Brownell

Bobby R. Alford Department of Otorhinolaryngology and Communicative Sciences, Baylor College of Medicine, Houston, Texas 77030

Alexander A. Spector

Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland 21205

(Received 28 July 2004; revised 11 January 2005; accepted 11 January 2005)

Cochlear outer hair cells (OHC) are critically important for the amplification and sharp frequency selectivity of the mammalian ear. The microchamber experiment has been an effective tool to analyze the OHC high-frequency performance. In this study, the OHC electrical stimulation in the microchamber is simulated. The model takes into account the inertial and viscous properties of fluids inside and outside the cell as well as the viscoelastic and piezoelectric properties of the cell composite membrane (wall). The closed ends of the cylindrical cell were considered as oscillatory rigid plates. The final solution was obtained in terms of Fourier series, and it was checked against the available results of the microchamber experiment. The conditions of the interaction between the cell and pipette was analyzed, and it was found that the amount of slip along the contact surface has a significant effect on the cell electromotile response. The cell's length changes were computed as a function of frequency, and their dependence on the viscosities of both fluids and the cell wall was analyzed. The distribution of the viscous losses inside the fluids was also estimated. The proposed approach can help in a better understanding of the high-frequency OHC electromotility under experimental and physiological conditions. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1863732]

PACS numbers: 43.64.Ld, 43.64.Kc, 43.64.Bt, 43.64.Nf [WS]

Pages: 2147–2157

I. INTRODUCTION

Cochlear outer hair cells (OHC) are crucial to the frequency discrimination and tuning during the mammalian hearing process (Geisler, 1998; Dallos, 1996; Ruggero, 1992). These cells are located between the underlying basilar and overlying tectorial membranes. It is believed that, via a novel electromotility (Brownell *et al.*, 1985, 2001), OHCs generate active forces that amplify and sharpen passive vibration of the basilar and tectorial membranes. Recently, a motor protein, named prestin, necessary for OHC electromotility and active hearing has been identified (Zheng *et al.*, 2000; Liberman *et al.*, 2002).

OHCs have shown their active properties in the whole acoustic frequency range. *In vitro* electrical stimulation of OHCs has demonstrated an electromotile response up to tens of kHz (Dallos and Evans, 1995; Frank *et al.*, 1999). It is probably even more important that the OHC is capable of generating a constant active force up to about 100 kHz. The transfer of an electric charge across the cell membrane is an intrinsic part of the mechanism of electromotility (Huang and Santos-Sacchi, 1994; Gale and Ashmore, 1997a). Gale and Ashmore (1997b) have shown that the cell has this ability at frequencies up to 40 kHz, and the amount of the translocated charge stays constant up to 20 kHz.

Despite important evidence for the high-frequency effectiveness of OHC, the mechanism of cells functioning under such conditions is not fully understood. One question is associated with the magnitude of the receptor potential that is supposed to drive the cell electromotile response. High-frequency membrane capacitive filtering of the receptor potential was discussed in the literature (Santo-Sacchi, 1992; Housley and Ashmore, 1992). A number of studies have addressed this problem by considering either a different mechanism of delivering the transmembrane potential (Dallos and Evans, 1995) or finer properties of the cell membrane (Santos-Sacchi *et al.*, 1998; Spector *et al.*, 2003, 2004; Weitzel *et al.*, 2003; Ospeck *et al.*, 2003).

Another key component of the high-frequency mechanics of the OHC is the interaction between the vibrating cell wall and the fluids inside and outside the cell. The intrinsic losses associated with the viscosity of the two fluids limit the high-frequency response of the cell. In a modeling study, Tolomeo and Steele (1998) have computed high-frequency cell's length changes in response to mechanical and electrical stimulations. The authors analyzed the effects of the extracellular and intracellular fluid viscosities considering them both equal to the viscosity of water. In addition to the mentioned factors, the viscosity of the cell wall also contributes to the overall viscous losses and can reduce high-frequency electromotility. This can be especially important for the OHC that has its cytoskeleton, the most viscous component, inside

^{a)}Electronic mail: zliao@bme.jhu.edu

the composite wall. Ratnanather *et al.* (1996) considered a viscoelastic cell wall and computed the high-frequency responses to a mechanical stimulation of the cell. While the estimation of the OHC's wall effective viscosity is still an open problem, there have been experimental studies that contribute to its solution. Ehrenstein and Iwasa (1996) have measured a time constant of the relaxation of tension on the OHC wall. Li *et al.* (2002), by pulling tethers from the membrane, have estimated the apparent membrane viscosity that included the viscous losses in the bilayer as well as friction of the movement of the membrane along the cell cytoskeleton.

The voltage-clamp experiment provides important information on the OHC's electromotile response and nonlinear capacitance. However, the resistive properties of the pipette clamping the membrane potential preclude the measurements under high-frequency conditions. This effect was overcome in the microchamber experiment where an electric field is applied across the cell membrane. Dallos *et al.* (1993), Evans *et al.* (1991), and Hallworth *et al.* (1993) have developed the microchamber technique and proposed a model of the OHC behavior under the condition of this experiment. Later, Dallos and Evans (1995) used this method to observe the high-frequency electromotile responses of OHC. By using the microchamber experiment, Frank *et al.* (1999) not only measured high-frequency electromotility but also estimated the high-frequency active force generated by the cell. Also, He and Dallos (2000) used the microchamber technique to measure the voltage-dependent stiffness of OHC.

The active and passive properties of the cell wall are critically important for the modeling of the OHC mechanics. The effective passive properties of the composite OHC wall can be described by a linear 2D viscoelastic model. The elastic part of the model is represented by the orthotropic elastic moduli (Tolomeo and Steele, 1995; Spector *et al.*, 1998; Iwasa and Adachi, 1997; Spector and Jean, 2003). Anisotropy of the cell wall is associated with the composition of the cell cytoskeleton made of two proteins with different properties. The viscous part of the model is described in terms of the wall viscosity (Ratnanather *et al.*, 1996). For small changes in the transmembrane potential, the active properties of the cell wall fit a linear piezoelectric model (Tolomeo and Steele, 1995; Spector, 2001; Spector and Jean, 2004). The OHC changes its dimensions in response to electrical stimulation, and conversely electric current can be generated across the cell membrane in response to a deformation of the cell. These properties constitute a piezoelectric-type behavior of the material.

In this study, we model the high-frequency electrical stimulation of OHC in the microchamber. In our model, the cell is considered as a circular cylinder with viscoelastic and piezoelectric lateral wall and closed ends. The solution to the problem is given in terms of Fourier series, similar to that proposed by Tolomeo and Steele (1998). We applied our model to the analysis of the data of the microchamber experiment (Evans *et al.*, 1991; Dallos *et al.*, 1993; Dallos and Evans, 1995; Frank *et al.*, 1999). We have shown that the parameters of the experiment, such as the amount of slip at the area of interaction between the cell and pipette and the

set-point potential, have significant effects on the cell response. This finding helps reconcile different versions of the microchamber experiment. We analyze the effects of the viscosities of the wall and of the two surrounding fluids. Finally, we compute the viscous losses inside the external and internal fluids as functions of the radial distance from the cell wall.

II. MODEL

A. Cell wall

We start with the constitutive relations for the cell wall; then, we consider the extracellular and intracellular fluids; after that we analyze the effect of the ends of the cell; and, finally, a solution to the problem is presented in terms of Fourier series. For the cell wall, we use a linear integrative model that describes both the passive and active properties of the OHC. The total resultant forces in the cell wall consist of three parts: orthotropic elastic, viscous, and piezoelectric. The first two of them are associated with the passive properties of the wall, and the third part is determined by its active properties. The equation for the resultants takes the form

$$\begin{bmatrix} N_x \\ N_\theta \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{12} & C_{22} \end{bmatrix} \begin{bmatrix} \frac{\partial u_x}{\partial x} \\ \frac{u_r}{r_c} \end{bmatrix} + \begin{bmatrix} \eta & -\eta \\ -\eta & \eta \end{bmatrix} \begin{bmatrix} \frac{\partial^2 u_x}{\partial x \partial t} \\ \frac{\partial u_r}{r_c \partial t} \end{bmatrix} + V \begin{bmatrix} e_x \\ e_\theta \end{bmatrix}, \quad (1)$$

where N_x and N_θ are the components of the stress resultant (i.e., the product of the stress and cell wall thickness) generated in the cell wall; the subscripts r , x , and θ indicate the radial, axial, and circumferential directions, respectively; C 's are the stiffness moduli; u_x and u_θ are two components of the displacement; η is the cell wall viscosity; t is the time; V is the transmembrane potential change; and e_x and e_θ are two coefficients that determine the production of the local active stress resultant per unit transmembrane potential (Spector *et al.*, 1998, 1999; Spector and Jean, 2003) (Fig. 1). The stress resultants are related to the cell wall tractions by the following equilibrium equations:

$$N_\theta = \sigma_r r_c, \quad (2)$$

$$N_x = - \int_a^x \sigma_{rx} d\bar{x} + N_{\text{end}}, \quad (3)$$

where σ is the traction acting on the cell wall surface, the subscripts r and rx indicate the radial and longitudinal directions, and N_{end} is the force produced by the end cap oscillating in the viscous fluid (see the discussion of this term below). The resultant cell wall tractions are determined from the extracellular and intracellular fluids, and other external sources

$$\vec{\sigma}_{\text{wall}} = \vec{\sigma}_{\text{sti}} + \vec{\sigma}_{\text{ext}} - \vec{\sigma}_{\text{int}}, \quad (4)$$

where $\vec{\sigma}_{\text{wall}}$ is the resultant cell wall traction in the vector form including two components σ_r and σ_{rx} ; $\vec{\sigma}_{\text{ext}}$ and $\vec{\sigma}_{\text{int}}$ are tractions on the cell wall surface exerted by the extracel-

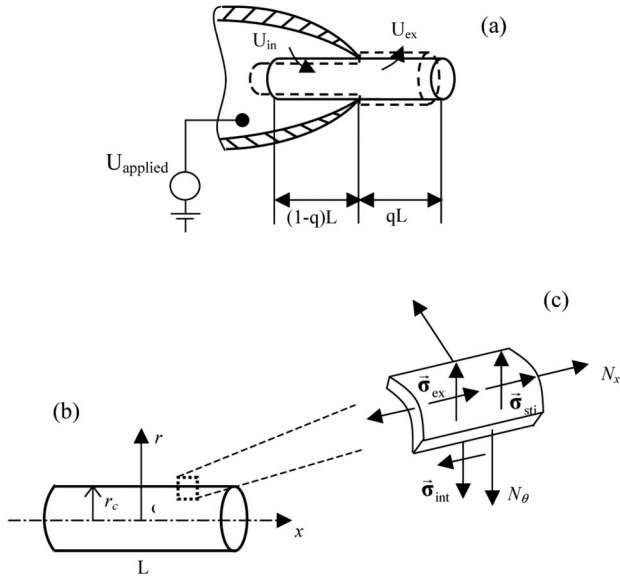


FIG. 1. OHC modeling in microchamber experiment. (a) Microchamber setup. U_{applied} is applied voltage. The transmembrane voltage change in the excluded and included parts are $U_{\text{ex}} = (1-q)U_{\text{applied}}$ and $U_{\text{in}} = -qU_{\text{applied}}$, respectively. Here, q is the exclusion ratio (excluded cell length divided by total cell length). (b) OHC is modeled as a thin-walled circular cylinder with closed ends. The cylindrical wall is surrounded by the viscous intracellular and extracellular fluids. (c) $\vec{\sigma}_{\text{ext}}$ and $\vec{\sigma}_{\text{int}}$ represent the tractions on the cell wall surface due to the extracellular and intracellular fluids, respectively. $\vec{\sigma}_{\text{sti}}$ indicates the applied mechanical stimulation on the cell wall surface from the external resource other than the extracellular and intracellular fluids. $\vec{\sigma}_{\text{ext}}$, $\vec{\sigma}_{\text{int}}$, and $\vec{\sigma}_{\text{sti}}$ are in the vector form, each with two components in the longitudinal and radial directions. These tractions and applied mechanical stimulation mentioned above are balanced by the stress resultant (i.e., the product of the stress and cell wall thickness) generated in the cell wall (N_x, N_θ).

lular and intracellular fluids, respectively; $\vec{\sigma}_{\text{sti}}$ is the applied mechanical stimulation on the cell wall surface from external sources other than extracellular and intracellular fluids. We include the term $\vec{\sigma}_{\text{sti}}$ to compare our results with those of

Tolomeo and Steele (1998), who considered a mechanical perturbation of the cell.

The force and displacement can be expressed as Fourier series defined in the cell domain. For a loading symmetric with respect to the cell center ($x=0$), only odd harmonics is needed, while an antisymmetric loading (such as that in the microchamber experiment) requires only even harmonics (Tolomeo, 1995; Tolomeo and Steele, 1998). Below, we present the equations corresponding to the antisymmetric case, and those for the symmetric case can be obtained by replacing the $\cos(\)$ functions with the $\sin(\)$ functions, and vice versa

$$u_x = \sum_{m=\text{even}}^{m \text{ max}} u_{mx} \cos(\beta_m x) e^{i\omega t},$$

$$u_r = \sum_{m=\text{even}}^{m \text{ max}} u_{mr} \sin(\beta_m x) e^{i\omega t}, \quad (5)$$

$$\sigma_{rx} = \sum_{m=\text{even}}^{m \text{ max}} \sigma_{mrx} \cos(\beta_m x) e^{i\omega t},$$

$$\sigma_r = \sum_{m=\text{even}}^{m \text{ max}} \sigma_{mr} \sin(\beta_m x) e^{i\omega t}, \quad (6)$$

$$Ve_x = \sum_{m=\text{even}}^{m \text{ max}} Q_m \sin(\beta_m x) e^{i\omega t},$$

$$Ve_\theta = \sum_{m=\text{even}}^{m \text{ max}} P_m \sin(\beta_m x) e^{i\omega t}, \quad (7)$$

$$N_{\text{end}} = \sum_{m=\text{even}}^{m \text{ max}} S_m \sin(\beta_m x) e^{i\omega t}, \quad (8)$$

where $\beta_m = m\pi/L$, $Q_m = Ve_x g_m$, $P_m = Ve_\theta g_m$, $S_m = N_{\text{end}} g_m$, and $g_m = -\{2[(-1)^{m/2} - 1]\}/(a\beta_m)$.

Substituting Eqs. (5)–(8) into Eqs. (2)–(3) and Eq. (1), we obtain the m th harmonic equation

$$\begin{bmatrix} \sigma_{mrx} \\ \sigma_{mr} \end{bmatrix} = \begin{bmatrix} (C_{11} + i\omega\eta)\beta_m^2 & -(C_{12} - i\omega\eta)\frac{\beta_m}{r_c} \\ -(C_{12} - i\omega\eta)\frac{\beta_m}{r_c} & (C_{22} + i\omega\eta)\frac{1}{r_c^2} \end{bmatrix} \begin{bmatrix} u_{mx} \\ u_{mr} \end{bmatrix} + \begin{bmatrix} -Q_m\beta_m \\ P_m/r_c \end{bmatrix} + \begin{bmatrix} S_m\beta_m \\ 0 \end{bmatrix}. \quad (9)$$

The above equation can be written in the matrix form

$$\vec{\sigma}_{m,\text{wall}} = \mathbf{k}_{m,\text{wall}} \vec{\mathbf{u}}_m + \vec{\sigma}_{m,\text{piez}} + \vec{\sigma}_{m,\text{end}}. \quad (10)$$

Finally, by combining all harmonics together, we obtain

$$\vec{\sigma}_{\text{wall}} = \mathbf{k}_{\text{wall}} \vec{\mathbf{u}}_{\text{wall}} + \vec{\sigma}_{\text{piez}} + \vec{\sigma}_{\text{end}}. \quad (11)$$

B. Intracellular and extracellular fluids

Fourier series technique is applied to solve linearized Navier–Stokes equations for the intracellular and extracellular fluids. In this method, the fluid velocity is at first decom-

posed into irrotational and rotational components, which are then expressed by the method of separation of variables using Bessel functions. The coefficients of Fourier series in the fluid harmonics are converted into the coefficients in the cell wall harmonics based on the no-slip boundary condition along the fluid–cell interface ($\vec{\mathbf{u}}_{\text{wall}} = \vec{\mathbf{u}}_{\text{fluid}}$). As a result of these derivations, the relationship between the fluid stress jump across cell wall ($\vec{\sigma}_{\text{ext}} - \vec{\sigma}_{\text{int}}$) and the cell deformation can be expressed by the following equation (Tolomeo, 1995; Tolomeo and Steele, 1998):

$$\vec{\sigma}_{\text{ext}} - \vec{\sigma}_{\text{int}} = -\mathbf{k}_{\text{fluid}} \vec{\mathbf{u}}_{\text{wall}}, \quad (12)$$

where $\mathbf{k}_{\text{fluid}}$ is the fluid stiffness matrix using cell wall harmonics.

C. Effect of closed ends

As OHC changes its length periodically, the end cap is subjected to a drag force produced by the viscous intracellular and extracellular fluids. If the end cap is assumed to be a rigid circular plate, the whole drag force will be transmitted to the cell wall. Zhang and Stone (1998) studied oscillatory motions of a circular disk in a viscous flow and gave a high-frequency asymptotic expression for the corresponding drag force

$$F_{\text{end}} = -16 \left[\frac{i\lambda^2}{6} + (1.32e^{i0.27\pi\lambda}) \right] \mu r_c(i\omega) u_{\text{end}} e^{i\omega t}, \quad (13)$$

where $\lambda^2 = \rho \omega r_c^2 / \mu$, μ is the fluid dynamic viscosity, u_{end} is the disk displacement, ω is the oscillatory frequency. We utilize this relationship and express the end force in terms of the end resultant N_{end}

$$N_{\text{end}} = \frac{F_{\text{end}}}{2\pi r_c} = \frac{k_{\text{end}}}{2\pi r_c} u_{\text{end}}, \quad (14)$$

where $k_{\text{end}} = -16[(i\lambda^2/6) + (1.32e^{i0.27\pi\lambda})] \mu r_c(i\omega) e^{i\omega t}$.

D. Final form

Substitution of Eqs. (11) and (12) into Eq. (4) yields

$$(\mathbf{k}_{\text{wall}} + \mathbf{k}_{\text{fluid}}) \vec{\mathbf{u}}_{\text{wall}} = \vec{\boldsymbol{\sigma}}_{\text{sti}} - \vec{\boldsymbol{\sigma}}_{\text{piez}} - \vec{\boldsymbol{\sigma}}_{\text{end}}. \quad (15)$$

Note that

$$\vec{\boldsymbol{\sigma}}_{\text{end}} = \frac{k_{\text{end}}}{2\pi r_c} \begin{pmatrix} g_2\beta_2 \\ 0 \\ g_4\beta_4 \\ 0 \\ g_6\beta_6 \\ 0 \\ \vdots \end{pmatrix} u_{\text{end}} = \frac{k_{\text{end}}}{2\pi r_c} \begin{pmatrix} g_2\beta_2 \\ 0 \\ g_4\beta_4 \\ 0 \\ g_6\beta_6 \\ 0 \\ \vdots \end{pmatrix} (\cos(\beta_2 a) \ 0 \ \cos(\beta_4 a) \ 0 \ \cos(\beta_6 a) \ 0 \ \cdots) \begin{pmatrix} u_{2x} \\ u_{2r} \\ u_{4x} \\ u_{4r} \\ u_{6x} \\ u_{6r} \\ \vdots \end{pmatrix} = \mathbf{k}_{\text{end}} \vec{\mathbf{u}}_{\text{wall}}. \quad (16)$$

With the above equation considered, Eq. (15) becomes

$$\vec{\mathbf{u}}_{\text{wall}} = [\mathbf{k}_{\text{wall}} + \mathbf{k}_{\text{fluid}} + \mathbf{k}_{\text{end}}]^{-1} (\vec{\boldsymbol{\sigma}}_{\text{sti}} - \vec{\boldsymbol{\sigma}}_{\text{piez}}). \quad (17)$$

Here, $\vec{\mathbf{u}}_{\text{wall}}$, $\vec{\boldsymbol{\sigma}}_{\text{sti}}$, and $\vec{\boldsymbol{\sigma}}_{\text{piez}}$ are, respectively, the vectors of the Fourier coefficients of the cell wall displacement, the traction acting on the cell wall surface due to the applied mechanical stimulation (other than those associated with the extracellular and intracellular fluids), and the stress due to the electrical stimulation of the cell. Also, \mathbf{k}_{wall} , $\mathbf{k}_{\text{fluid}}$, and \mathbf{k}_{end} are the matrices that determine the stiffness associated with the cell wall, fluids, and closed end, respectively. Following Tolomeo and Steele (1998), we use 10 harmonics in the cell wall domain and 20 in the fluid domain.

The general procedure of obtaining the cell displacement is described as follows: at first, the applied mechanical or electrical stimulation is expressed in terms of Fourier series; then, this coefficient vector is substituted into Eq. (17) to calculate the vector consisting of coefficients of Fourier series for the displacement; and finally, the cell displacement is calculated from its Fourier series coefficients.

III. RESULTS

A. Parameters used

The parameters of our model include the cell wall orthotropic moduli, the coefficients e_x and e_θ that determine the cell active properties, and the viscosities of the fluids and the cell wall. Spector *et al.* (1998) have derived the relationships between the elastic moduli of the cell wall and the axial stiffness of the OHC, and Spector and Jean (2003) have shown that the results obtained on the basis of different sets

of experiments (Iwasa and Adachi, 1997; Spector *et al.*, 1998) converge if the cell stiffness is about 0.1 N/m. Such cell stiffness is consistent with the latest measurements (Hallworth, 1997). This results in the following estimates of the elastic moduli that we use in our modeling: $C_{11} = 0.096$ N/m, $C_{12} = 0.16$ N/m, $C_{22} = 0.3$ N/m. The coefficients e_x and e_θ are determined by both the elastic moduli (Tolomeo and Steele, 1995; Spector *et al.*, 1999) and the active strains extracted from Dallos *et al.* (1993). The resulting coefficients are equal to $e_x = 0.0029$ N/Vm, $e_\theta = 0.0018$ N/Vm. The viscous properties of the extracellular fluid are similar to those of water, but the viscosity of the intracellular fluid is probably different (higher) because of the presence of various proteins (e.g., G-actin). Fung (1993) mentioned the internal viscosity in some cells being 5–6 times greater than that of water. Luby-Phelps (2000) reported a range of 1×10^{-3} – 10×10^{-3} Ns/m² for cytoplasmic viscosity. While the OHC intracellular viscosity has not yet been measured, we assume that the OHC cytoplasmic viscosity is within a range 1×10^{-3} – 10×10^{-3} Ns/m². The viscosity of the OHC wall is another parameter that has not been measured experimentally. Assuming that the cytoskeleton is the most viscous component of the composite wall, and also that the longitudinally directed spectrin molecules are mainly responsible for the viscous losses in the vibrating cell, we chose the cell wall viscosity on the order of magnitude of 1.0×10^{-7} Ns/m (see further discussion of this parameter below); note that this parameter is expressed as cell surface viscosity and thus has a different dimension than the cytoplasmic viscosity.

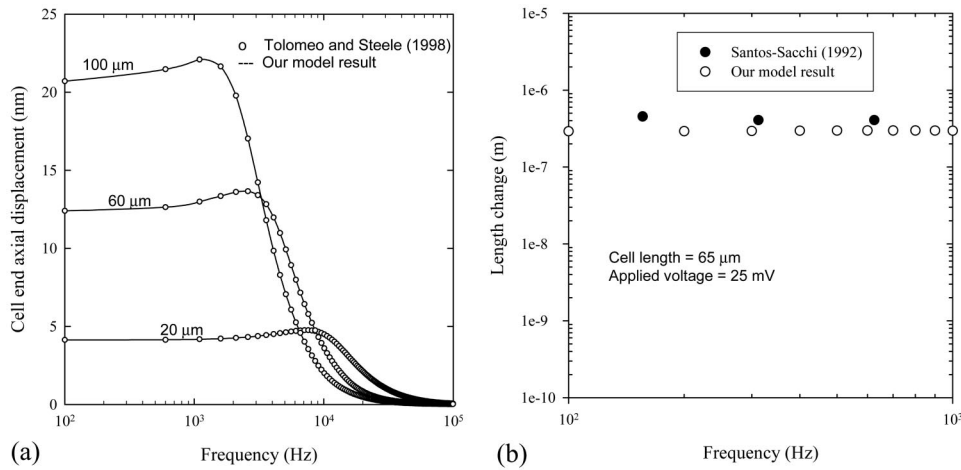


FIG. 2. Validation of our model. (a) Tolomeo and Steele's model (1998) is a specific case that can be achieved by our model through nullifying the effect of the cell wall viscosity and closed end. For the purpose of verifying our code, an example from Tolomeo and Steele's paper (1998) is used to compare with our model results. Our model reproduces Tolomeo and Steele's results. Cell lengths are 20, 60, and 100 μm , and 1-Pa radial stress in the form of first harmonic is applied to the cell wall. (b) Comparison with Santos-Sacchi's experiment (1992). They are in good agreement in the low frequency (less than 1 kHz).

B. Validation of the model

The first validation case is the comparison of the results of our model with Tolomeo and Steele's (1998), who considered OHC as a cylinder with open ends and a purely elastic wall. Thus, we let $\mathbf{k}_{\text{end}}=0$ and $\eta=0$ in Eq. (17) and use the same elastic moduli as Tolomeo and Steele (1998). The externally applied mechanical stimulation is prescribed at 1-Pa radial pressure (stress) in the form of the first harmonic. The loading is symmetric with respect to the cell center ($x=0$). As shown in Fig. 2(a), our model reproduces the results of Tolomeo and Steele (1998).

We use the results of the voltage-clamp experiment (Santos-Sacchi, 1992) as a second validation case. To compare our results with those of Santos-Sacchi (1992), we used the version of the proposed model that is based on no-slip condition along the cell/microchamber interface (see further discussion below). Under this condition, each excluded or included part of the cell will be under conditions equivalent to those in the voltage-clamp experiment if the cell lengths and the transmembrane voltages are the same. In the comparison to Santos-Sacchi's (1992) experiment, we use the elastic moduli and ranges of the fluid and cell wall viscosities discussed in the previous section. However, the viscosity is probably not important in this case because the experimental frequencies are limited to about 1 kHz. The results of the comparison are presented in Fig. 2(b), showing that our predictions are in good agreement with the experimental data in the low frequency (less than 1 kHz). In the high frequency, Santos-Sacchi's experimental data are complicated by the limitation of his instrumentation.

C. Modeling microchamber experiment

In the microchamber experiment, the cell is held by the micropipette along a part of its lateral wall. Below, we show that the amount of slip along the cell/micropipette interface is a significant parameter of the microchamber experiment, and the differences in the interface slip conditions can explain discrepancies in different versions of this experiment. To analyze the effect of the interface slip, we consider two extreme cases: no-slip and full-slip (no-friction) conditions. It is expected that the results of the general case with some

slip and friction will fall within the range between these two extreme cases. For low frequencies, Fig. 3(a) shows that the no-slip condition results in an electromotile response two times greater than that corresponding to the full-slip condition, but this difference reduces at high frequencies. Also, there is a difference in the phase shifts corresponding to these two conditions, especially at high frequencies [Fig. 3(b)]. The axial displacement profiles along the cell are provided in Fig. 4. They show that the zero-displacement point is located at the center of the whole cell length for the no-slip condition and at the center of the excluded/included part of the cell for the full-slip condition, thus causing different effective lengths and electromotile responses in these two cases. Frank *et al.* (1999) obtained smaller cell length changes compared to those by Dallos and Evans (1995), and the reason for that may come from the fact that the set point in Frank's *et al.* experiment (1999) is not at the steepest part of OHC electromechanical transduction curve. Figure 5 shows that indeed by shifting set-point potential to more hyperpolarization our model results can be made closer to those in Frank's *et al.* experiment (1999) (see also the discussion below).

We assume that the ratio of the values of $e_x(e_\theta)$ estimated at two points is equal to that of the corresponding slopes of the electromechanical transduction curve. As a result, it is found that the set-point potential in Frank's experiment may be a -100 mV shift from the normal resting potential or more (Fig. 5).

In Fig. 6, we show the effect of closed ends of the cell on the magnitude and phase of the cell displacement. Although the difference in the magnitude between the closed and open ends is small, the difference in phase is relatively more significant, especially at high frequencies. Figure 7 indicates that an increase in the cell wall viscosity removes the peak in the cell displacement and reduces the corner frequency (i.e., the frequency at which the amplitude of displacement starts to decrease). It is shown in Fig. 8 that an increase in the intracellular fluid viscosity also reduces the corner frequency.

The fluid viscous damping is important for the OHC high-frequency performance. The dissipative energy Φ per volume per time is defined by the following equation (Currie, 1974):

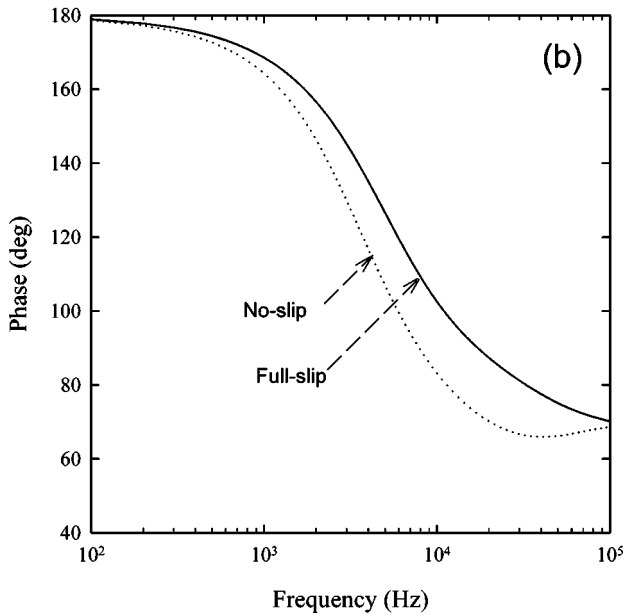
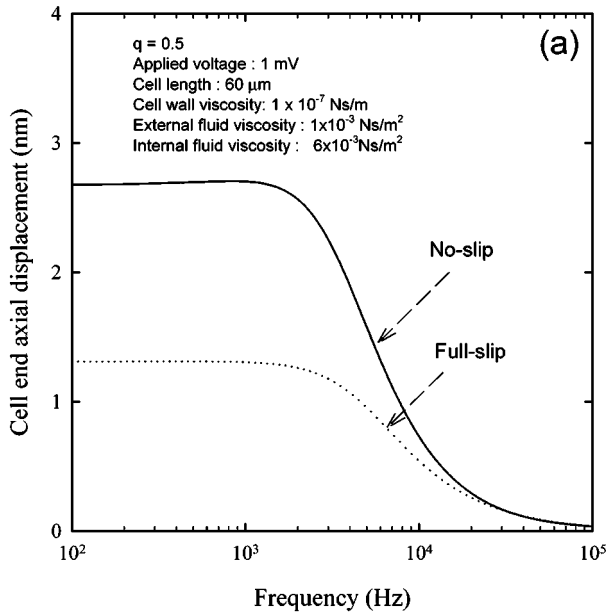


FIG. 3. Modeling microchamber experiment under the full-slip and no-slip conditions (a) amplitude (b) phase. The cell is half included and half excluded.

$$\Phi = 2\mu \left[\left(\frac{\partial v_r}{\partial r} \right)^2 + \left(\frac{v_r}{r} \right)^2 + \left(\frac{\partial v_x}{\partial x} \right)^2 \right] + \mu \left(\frac{\partial v_x}{\partial r} + \frac{\partial v_r}{\partial x} \right)^2, \quad (18)$$

where μ is dynamic viscosity of the fluid, and v_r and v_x are radial and axial velocities, respectively. Figure 9 displays the distribution of energy loss along the radial direction under different frequencies 5, 10, and 50 kHz. The energy loss profiles behave like a classic Stokes boundary layer near an oscillating plate (Lu, 1977), which decrease exponentially away from the cell wall.

Figure 10 shows the results for different exclusion ratio (q) in the microchamber setup. Here, $q = 0.5$ yields the maximum displacement due to the maximum production of excluded cell length and membrane potential at that cell part.

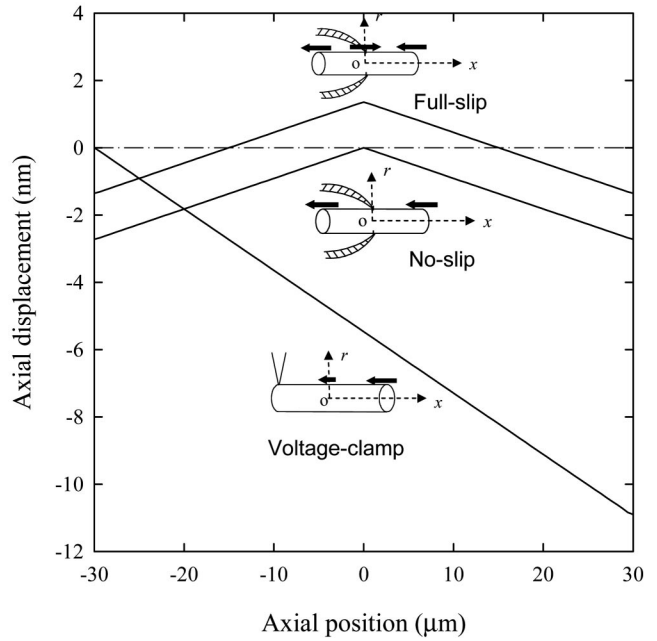


FIG. 4. The axial displacement profiles along the cell. The effective length, which determines cell end movement, is different in voltage-clamp and microchamber (full-slip/no-slip) setup.

IV. DISCUSSION

The amount of slip along the cell/micropipette interface has not been reported in earlier papers (Dallos *et al.*, 1993; Dallos and Evans, 1995; Frank *et al.*, 1999). However, recently, He and Dallos (2000) suggested that the interface slip should be small. It is still important to investigate the sensitivity of the cell electromotile response to the boundary condition along the cell–micropipette interface. Figure 3 shows that the length change under the no-slip condition is larger than that corresponding to the full-slip condition. To explore the underlying mechanism, the distribution of the axial dis-

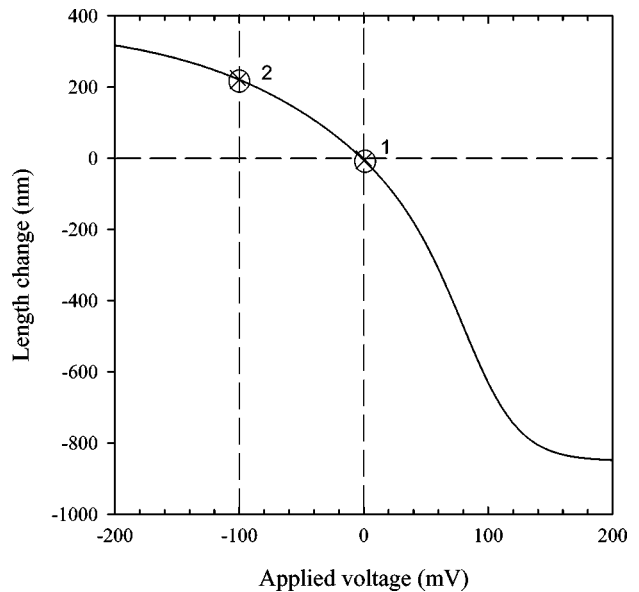


FIG. 5. Estimation of set-point in Frank's *et al.* (1999) experiment. The electromechanical transduction function curve comes from Dallos *et al.* (1993). Two points are selected to calculate electromechanical responses.

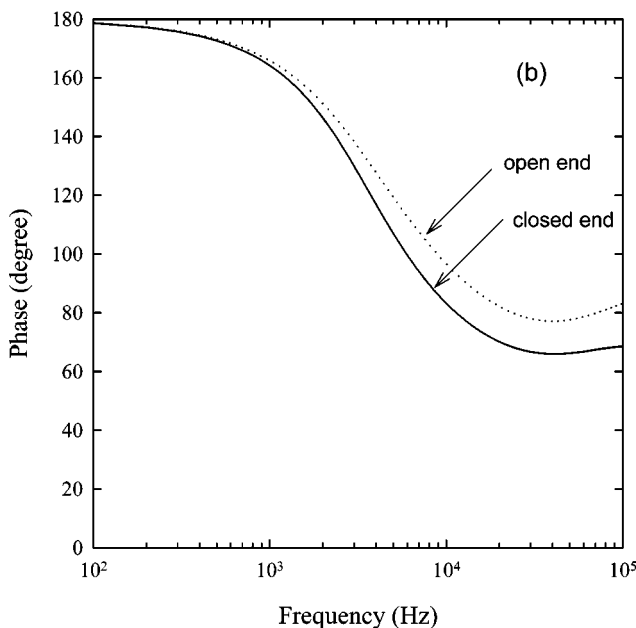
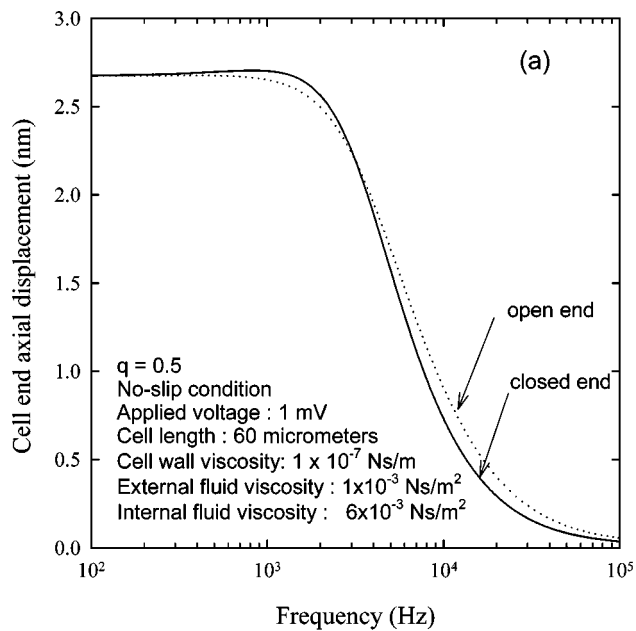


FIG. 6. Closed-end and open-end cases (a) amplitude (b) phase. The difference in the amplitude is small, but the difference in the phase shift reaches around 15 deg at high frequency.

placements along the cell length is given in Fig. 4. Under the no-slip condition, the holding point with the zero displacement is at the cell center, between the excluded and included parts. In this case, the effective length is equal to a half of a cell length. Under the full-slip condition, due to the cell sliding at the interface, the zero-displacement point moves to the center of the excluded–included part of the cell. In contrast to the case of the no-slip condition, the effective length reduces to one-quarter of a cell length, resulting in smaller cell motility, which is half of that corresponding to the no-slip condition. Figure 3 also shows how the corner frequency is affected by the boundary condition with the greater value corresponding to the full-slip condition. In addition, Fig. 3 shows the phase shift versus frequency for two types of

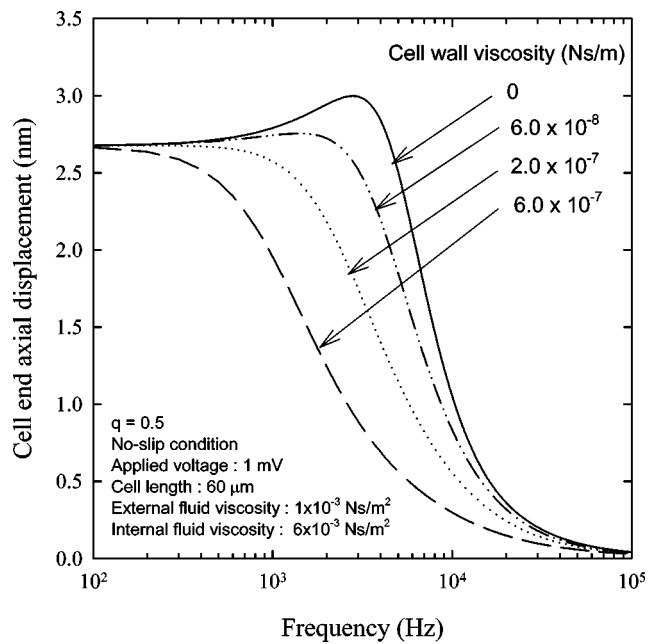


FIG. 7. Effect of the cell wall viscosity (η) on cell electromotility.

boundary conditions. The low-frequency shift is close to 180 deg, which corresponds to a contraction of the cell in response to its depolarization.

Evans *et al.* (1991) and Dallos and Evans (1993) reported electromotility range 0.92–4.74 nm/mV for the cells of 55–72 μm in length when the cell is half-included in the microchamber. These experimental results are in good agreement with our model predictions (1.3–2.7 nm/mV) for the cell length 60 μm . Evans *et al.* (1991) acknowledged that the OHC electromotility may vary significantly (0.92–4.74 nm/mV) even for very similar cells (55–72 μm), and they pointed out potential reasons for that, including individual cell’s variability and possible damage due to isolation. Another cause for this effect, suggested by our model, is the

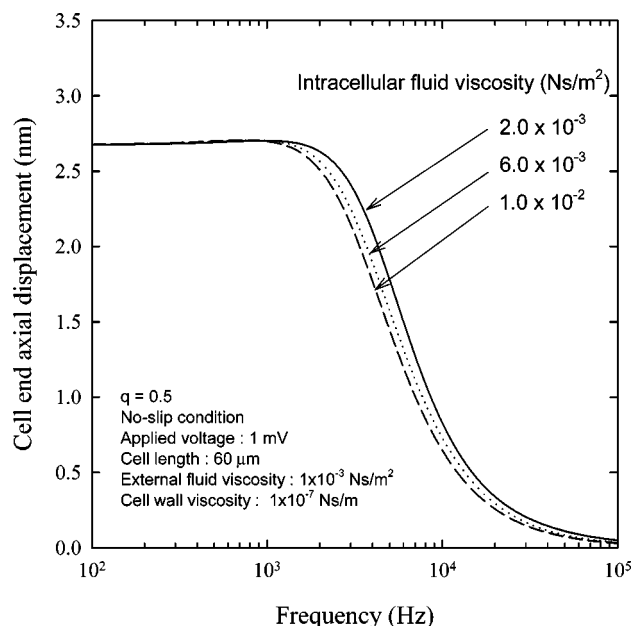


FIG. 8. Effect of the intracellular fluid viscosity on cell electromotility.

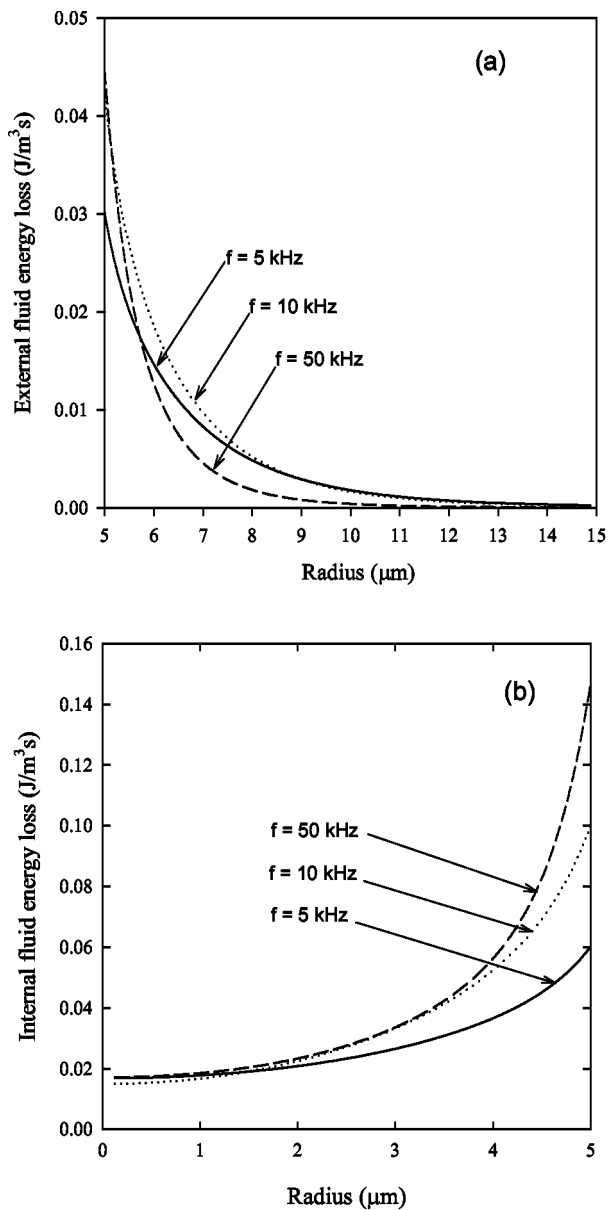


FIG. 9. Distributions of the volume density of the energy loss along radial direction in the (a) external fluid and (b) internal fluid. The axial position is $15 \mu\text{m}$ away from cell center. The frequency is chosen at 5, 10, and 50 kHz. The energy loss in the external fluid decreases exponentially away from the cell wall, indicating a boundary layer effect.

holding condition at the cell–pipette interface. The holding could be loose and the cell could move in and out of the chamber, resulting in a reduction of the electromotile response. Our model predicts that this reduction can reach up to 50%.

In the experiment of Frank *et al.* (1999), the observed cell’s end displacement amplitude (parameter $x_{[\text{nm}]}$ in their paper) is 0.25 nm/mV and the lengths of the cell and its excluded part are 51 and $24 \mu\text{m}$, respectively. The prediction of our model for the cell of a comparable size (the lengths of the cell and its excluded part are 51 and $25.5 \mu\text{m}$, respectively) is in the range of 1.14 – 2.28 nm/mV . These values correspond to the resting potential of -70 mV , and the lower and upper limits come from the full- and no-slip conditions, respectively. Frank *et al.* (1999) noted that their value is

much smaller than that in other experiments (Dallos and Evans, 1995) and pointed out that perhaps the set point in their experiments was not located at the steepest part of the OHC electromechanical transfer function curve. In order to estimate the effect of the set point, we recalculated our predictions assuming more negative (farther from the steepest point of the electromotile response) resting potentials. By assuming the coefficient e_x or e_θ is proportional to the slope of the electromechanical transduction curve, we found that, if the set-point potential shifts 100 mV toward higher polarization, the predicted cell displacement amplitude reduces to 0.25 – 0.5 nm/mV depending on the chosen boundary condition at the cell–pipette interface. Therefore, the results of our modeling are consistent with the available data from the microchamber experiment (Evans *et al.*, 1991, Dallos and Evans, 1995, Frank *et al.*, 1999), and the measurements of different groups can be reconciled by taking into account the set point of the electric potential and the boundary conditions at the cell–pipette interface.

For the cell of $60 \mu\text{m}$ in length, the estimated corner frequency is around 1 – 2 kHz where the lower and upper limits correspond to the no- and full-slip conditions, respectively. For cells of comparable lengths, the corner frequency in Frank’s *et al.* (1999) experiment is about 20 – 30 kHz . Frank *et al.* (1999) have defined $f_{3 \text{ dB}}$ as the frequency at which the amplitude drops 3 dB (50%) from its low-frequency value. Our estimate of $f_{3 \text{ dB}}$ is in the range 6 – 10 kHz depending on the boundary condition. Qualitatively, the computed electromotile response as a function of frequency is similar to that observed by Frank *et al.* (1999). Dallos and Evans (1995) have reported the outer hair cell electromotile response as a two-plateau function of the frequency. Our results do not show the second (high-frequency) plateau of Dallos and Evans’s (1995) experiment. The differences between the results of these two versions of the microchamber experiment could be explained by differences in experimental setups, including the cell’s slip conditions, cell stiffness, etc., but this issue has to be investigated further.

Figure 6 indicates that the effect of the closed ends on the displacement amplitude is small, but it becomes more significant for the phase shift. The difference between two phase shifts increases with frequency and it reaches around 10 deg at 10 kHz and 15 deg at 100 kHz . The results in Fig. 6 correspond to the no-slip conditions, but the effect for the full-slip condition (not shown) is similar.

Our analysis of the viscous effects (Figs. 7 and 8) shows that they result in faster decays in the electromotile responses of the cell and smaller values of the corner frequency. The viscosities of the fluids and the cell wall decrease peaks in the amplitude near the corner frequency, and for realistic values of the parameters, no peaks were found. This is consistent with the results of the microchamber experiment where no tuning was observed in the electromotile response of OHC. Figure 7 demonstrates the effect of the wall viscosity on the peak in the cellular response: increase in the wall viscosity results in the complete elimination of the peak in the cell displacement. Tolomeo (1995) has shown a similar effect of the viscosity of the intracellular fluid. We consider the effect of this viscosity in Fig. 8. In our case, its effect is

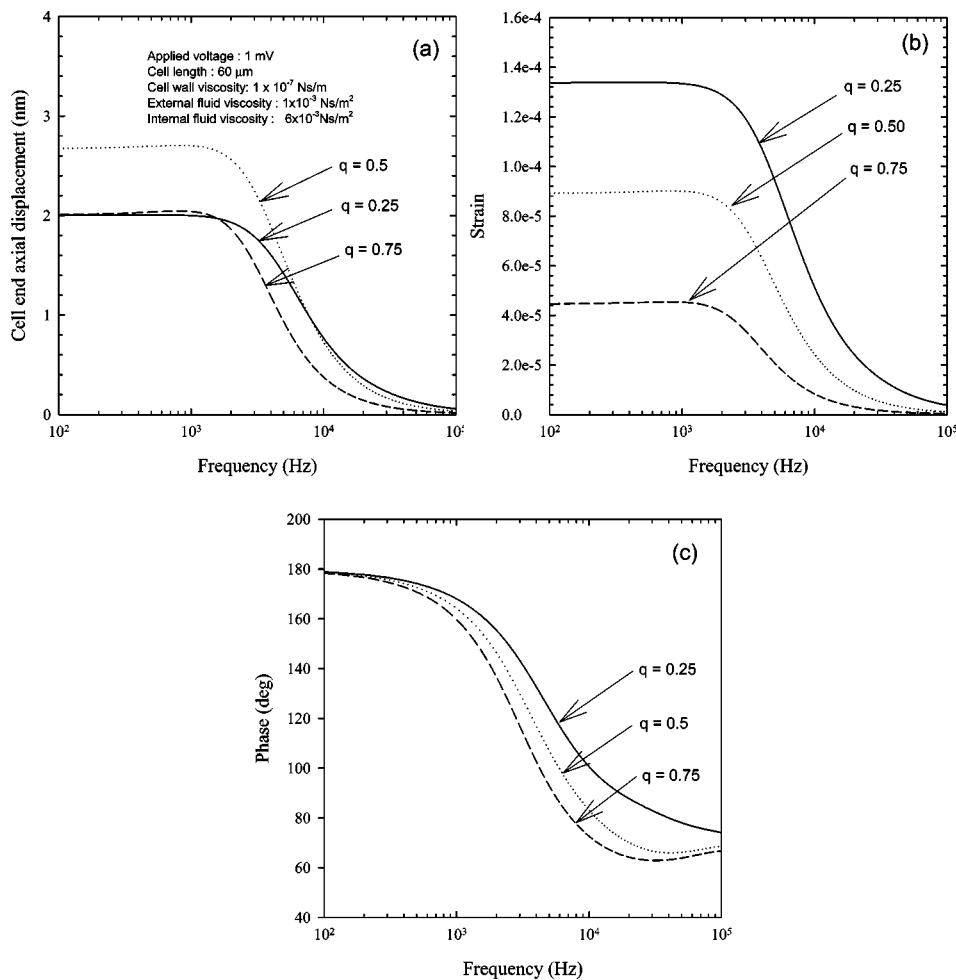


FIG. 10. Effect of exclusion ratio (q) in the microchamber setup (a) amplitude, (b) strain (cell end displacement divided by the length of the excluded part), and (c) phase shift. No-slip condition is considered.

smaller than that in Tolomeo's (1995) study because of greater values of additional viscosity of the cell wall.

We chose the viscosity of the external fluid equal to that of water, and we considered a range of the viscosities of the internal fluid greater than the water viscosity. The cell wall is a composite whose main components are the subsurface cisternae, cortical lattice (cytoskeleton), and the plasma membrane. There is a layer of the extracisternal fluid between the cytoskeleton and the plasma membrane. The bulk viscosities of the subsurface cisternae and the plasma membrane are probably close to that of the lipid bilayer, and the viscosity of the extracisternal fluid is probably equal to that of the intracellular fluid. As in the red blood cell that also has a composite membrane, the most viscous component of the cell wall is probably its cytoskeleton. We assumed that the cytoskeletal spectrin molecules oriented approximately parallel to the cell axis determine the viscous losses associated with the longitudinal vibration of the cell in the microchamber. The viscosity of the spectrin network in red blood cell has been estimated (Evans and Skalak, 1980). The topology of the spectrin networks in these two cells (OHC and red blood cell) is different, but the distances between two neighboring spectrin molecules are close to each other. On the basis of these arguments, we chose the viscosity of the OHC wall close to that of the red blood cell and its cytoskeleton.

The viscous loss is the major contributor to the electromotility decay with frequency, although losses of other na-

ture (e.g., electrical losses in the cell membrane) also play a role. Figure 9 shows the pattern of the distribution of losses (volume densities) within two fluids. The greatest loss occurs at the cell wall, followed by a steep decrease with the distance from the wall. This pattern constitutes a viscous boundary layer, which is more pronounced in the external fluid. We checked our computation of energy loss inside the fluids against the classic Stokes' problem on a vibrating rigid plate in the viscous fluid (Currie, 1974; Lu, 1977), and found a qualitative similarity in these two solutions.

The viscosities of the intracellular fluid and the cell wall have not been measured experimentally, and we have shown the significance of these parameters to high-frequency electromotility of OHC. Thus, our results can stimulate further experiments aimed at the measurements of the viscous parameters of this cell and its components.

In addition to the cell half-included/excluded condition ($q=0.5$) discussed above, our model can be applied to other situations of different exclusion ratio q (Fig. 10). In Fig. 10(a), $q=0.5$ yields the maximum electromotile response under the 1 mV applied voltage, because of the maximum production of the effective length and membrane potential at that cell part. This result is consistent with the result of Dallos *et al.* (1993) in the case of small voltages. In the case when, instead of the absolute length change, the strain (length change divided by the length of the excluded part of the cell) is used [Fig. 10(b)], the effect of parameter q re-

duces to the effect of the transmembrane potential. The transmembrane potential is a decreasing function of q . Similar pattern is observed in the phase shift [Fig. 10(c)].

In the present paper, we concentrate on the frequency effects associated with the interaction of the cell wall with the surrounding viscous fluids and on the viscous effects in the cell wall itself. We also consider the cell as finite, and take into account the distributions of the mechanical characteristics (stresses and strains) along the cell. However, we assume the transmembrane potential and the piezoelectric properties to be uniform. There have been reports on the nonuniform distribution of ion channels (Santos-Sacchi *et al.*, 1997) as well as heterogeneity of the molecular motors (Santos-Sacchi, 2002) along the cell that may have an effect on the piezoelectric and electrical properties of the cell. These effects will be taken into account in the later versions of our model.

Although the corner frequency of the electromotile displacement computed here is lower than the upper limit of the acoustic range, our ultimate interest is in the frequency range of the active force, the factor that determines the amplification of the basilar-membrane vibration. The corner frequency of the active force is greater than that of displacement of a free cell. This is because the cell has to be constrained in order to produce an active force. Such conditions correspond to smaller movement of the cell resulting in smaller viscous losses and higher corner frequency (the analysis of the high-frequency active force production will be published elsewhere).

The ultimate goal of our modeling is to better understand the cochlear amplifier and outer hair cell electromotility under high-frequency conditions. An advanced model of the cochlea that includes “a module,” where the electromechanical properties of an individual outer hair cell are included, can be an effective tool to achieve this goal. A logical first step to develop such a module would be to model the high-frequency performance of an unconstrained outer hair cell. Further steps would be to incorporate the interaction of the cell with the basilar and tectorial membrane, respectively, underlying and overlying the cell. Finally, the module has to be incorporated into the model of the whole cochlea that takes into account the hydrodynamic interaction and anatomical gradients along the cochlea. Thus, this paper presents the results of the first step toward advanced cochlear models. The next step associated with the high-frequency performance of the constrained cell, including the active force production, will be published elsewhere.

In this paper, we modeled high-frequency electromotility of outer hair cells electrically stimulated in the microchamber. We found that the parameters of the microchamber experiment, such as the boundary conditions at the cell–pipette interface and the set-point potential, have a significant effect on the cell response observed in the experiment, and this finding helps reconcile the differences in the existing measurements. The viscous effects (particularly that of the cell wall) eliminate peaks in the electromotile response and play the main role in its decay with frequency. We took into account the cell closed-end effects, and found them to be not very significant. Finally, we identified parameters of the cell

(the viscosities of the wall and intracellular fluid) important for the cell high-frequency performance that have not been experimentally measured in the OHC, and our results should provide a stimulus for conducting new experiments. The developed approach can be used in advanced cochlear models aimed at the analysis of the active processes accompanying mammalian hearing.

ACKNOWLEDGMENTS

This work was supported by research grants DC02775 and DC00354 from the National Institute of Deafness and other Communication Disorders (NIH).

- Brownell, W. E., Bader, C. D., Bertrand, D., and de Ribaupierre, Y. (1985). “Evoked mechanical responses of isolated cochlear outer hair cells,” *Science* **224**, 194–196.
- Brownell, W. E., Spector, A. A., Raphael, R. M., and Popel, A. S. (2001). “Micro- and nanomechanics of the cochlear outer hair cell,” *Annu. Rev. Biomed. Eng.* **3**, 169–194.
- Currie, C. E. (1974). *Fundamental Mechanics of Fluids* (McGraw-Hill, New York).
- Dallos, P. (1996). “Overview: Cochlear Neurobiology,” in *The Cochlea*, edited by P. Dallos, A. N. Popper, and R. R. Fay (Springer, New York), pp. 1–43.
- Dallos, P., and Evans, B. N. (1995). “High frequency motility of outer hair cells and the cochlear amplifier,” *Science* **267**, 2006–2009.
- Dallos, P., Hallworth, R., and Evans, B. N. (1993). “Theory of electrically driven shape changes of cochlear outer hair cells,” *J. Neurophysiol.* **70**, 299–323.
- Ehrenstein, D., and Iwasa, K. H. (1996). “Viscoelastic relaxation in the membrane of the auditory outer hair cell,” *Biophys. J.* **71**, 1087–1094.
- Evans, B. N., Hallworth, R., and Dallos, P. (1991). “Outer hair cell electromotility: The sensitivity and vulnerability of the dc component,” *Hear. Res.* **52**, 288–304.
- Evans, E. A., and Skalak, R. (1980). *Mechanics and Thermodynamics of Biomembranes* (CRC Press, Boca Raton, FL).
- Frank, G., Hemmer, W., and Gummer, A. W. (1999). “Limiting dynamics of high-frequency electromechanical transduction of outer hair cells,” *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4420–4425.
- Fung, Y. C. (1993). *Biomechanics: Mechanical Properties of Living Tissues* (Springer, New York).
- Gale, J. E., and Ashmore, J. F. (1997a). “The outer hair cell motor in membrane patches,” *Pfluegers Arch.* **434**, 267–271.
- Gale, J. E., and Ashmore, J. F. (1997b). “An intrinsic frequency limit to the cochlear amplifier,” *Nature (London)* **389**, 63–66.
- Geisler, C. D. (1998). *From Sound to Synapse* (Oxford University Press, New York).
- Hallworth, R. (1997). “Modulation of OHC force generation and stiffness by agents known to affect hearing,” in *Diversity in Auditory Mechanics*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, and E. Hecht-Poinar (World Scientific, Singapore), pp. 524–530.
- Hallworth, R., Evans, B. N., and Dallos, P. (1993). “The location and mechanism of electromotility in guinea pig outer hair cells,” *J. Neurophysiol.* **70**, 549–58.
- He, D. Z. Z., and Dallos, P. (2000). “Properties of voltage-dependent somatic stiffness of cochlear outer hair cells,” *J. Assoc. Res. Otolaryngol.* **1**, 64–81.
- Housley, G. D., and Ashmore, J. F. (1992). “Ionic currents of outer hair cells isolated from the guinea-pig cochlea,” *J. Physiol. (London)* **448**, 73–98.
- Huang, G., and Santos-Sacchi, J. (1994). “Motility voltage sensor of the outer hair cell resides within the lateral plasma membrane,” *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12268–12272.
- Iwasa, K. H., and Adachi, M. (1997). “Force generation in the outer hair cell of the cochlea,” *Biophys. J.* **73**, 546–555.
- Li, Z., Anvari, B., Takashima, M., Brecht, P., Torres, J. H., and Brownell, W. E. (2002). “Membrane tether formation from outer hair cells with optical tweezers,” *Biophys. J.* **82**, 1386–1395.
- Liberman, M. C., Gao, J., He, D. Z., Wu, X., Jia, S., and Zuo, J. (2002). “Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier,” *Nature* **419**, 300–304.

- Lu, P. (1977). *Introduction to the Mechanics of Viscous Fluids* (Hemisphere, Washington).
- Luby-Phelps, K. (2000). "Cytoarchitecture and physical properties of cytoplasm: Volume, viscosity, diffusion, intracellular surface area," *Int. Rev. Cytol.* **192**, 189–221.
- Ospeck, M., Dong, X. X., and Iwasa, K. H. (2003). "Limiting frequency of the cochlear amplifier based on electromotility of outer hair cells," *Biophys. J.* **84**, 739–749.
- Ratnanather, J. T., Spector, A. A., Popel, A. S., and Brownell, W. E. (1996). "Is the outer hair cell wall viscoelastic?," in *Proceedings of the Congress in Diversity in Auditory Mechanics*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, and E. Hecht-Poinar (World Scientific, Singapore), pp. 601–607.
- Ruggero, M. A. (1992). "Responses to sound of the basilar membrane of the mammalian cochlea," *Curr. Opin. Neurobiol.* **2**, 449–456.
- Santos-Sacchi, J. (1992). "On the frequency limit and phase of outer hair cell motility: Effects of the membrane filter," *J. Neurosci.* **12**, 1906–1916.
- Santos-Sacchi, J. (2002). "Functional motor microdomains of the outer hair cell lateral membrane," *Pfluegers Arch.* **445**, 331–336.
- Santos-Sacchi, J., Huang, G. J., and Wu, M. (1997). "Mapping the distribution of outer hair cell voltage-dependent conductances by electrical amputation," *Biophys. J.* **73**, 1424–1429.
- Santos-Sacchi, J., Kakehata, S., Kikuchi, T., Katori, Y., and Takasaka, T. (1998). "Density of motility-related charge in the outer hair cell of the guinea pig is inversely related to best frequency," *Neurosci. Lett.* **256**, 155–158.
- Spector, A. A. (2001). "A nonlinear electroelastic model of the auditory outer hair cell," *Int. J. Solids Struct.* **38**, 2115–2129.
- Spector, A. A., and Jean, R. P. (2003). "Elastic moduli of the piezoelectric cochlear outer hair cell membrane," *Exp. Mech.* **43**, 355–360.
- Spector, A. A., and Jean, R. P. (2004). "Modes and balance of energy in the piezoelectric cochlear outer hair cell wall," *J. Biomech. Eng.* **126**, 17–25.
- Spector, A. A., Brownell, W. E., and Popel, A. S. (1998). "Estimation of elastic moduli and bending stiffness of the anisotropic outer hair cell wall," *J. Acoust. Soc. Am.* **103**, 1007–1011.
- Spector, A. A., Brownell, W. E., and Popel, A. S. (1999). "Nonlinear active force generation by cochlear outer hair cell," *J. Acoust. Soc. Am.* **105**, 2414–2420.
- Spector, A. A., Brownell, W. E., and Popel, A. S. (2003). "Effect of outer hair cell piezoelectricity on high-frequency receptor potentials," *J. Acoust. Soc. Am.* **113**, 453–461.
- Spector, A. A., Popel, A. S., Eatock, R. A., and Brownell, W. E. (2004). "The effect of mechrosensitive channels in the lateral wall on high- and moderate frequency receptor potentials of the outer hair cell," in "Abstract of the 27th Annual Midwinter Research Meeting of ARO," February, 2004, Daytona Beach, FL, p. 108.
- Tolomeo, J. A. (1995). "Models of the structure and motility of the auditory outer hair cell," Ph.D. thesis, Stanford University.
- Tolomeo, J. A., and Steele, C. R. (1995). "Orthotropic piezoelectric properties of cochlear outer hair cell wall," *J. Acoust. Soc. Am.* **97**, 3006–3011.
- Tolomeo, J. A., and Steele, C. R. (1998). "A dynamic model of outer hair cell motility including intracellular and extracellular viscosity," *J. Acoust. Soc. Am.* **103**, 524–534.
- Weitzel, E. K., Tasker, R., and Brownell, W. E. (2003). "Outer hair cell piezoelectricity: Frequency response enhancement and resonance behavior," *J. Acoust. Soc. Am.* **114**, 1462–1466.
- Zhang, W., and Stone, H. A. (1998). "Oscillatory motions of circular disks and nearly spherical particles in viscous flows," *J. Fluid Mech.* **367**, 329–358.
- Zheng, J., Shen, W., He, D. Z.-Z., Long, K. B., Madison, L. D., and Dallos, P. (2000). "Prestin is the motor protein of cochlear outer hair cell," *Nature (London)* **405**, 149–155.

Modulation masking produced by second-order modulators^{a)}

Christian Füllgrabe^{b)}

Laboratoire de Psychologie Expérimentale—UMR CNRS 8581, Université René Descartes—Paris 5,
71 avenue Edouard Vaillant, 92774 Boulogne-Billancourt, France

Brian C. J. Moore

Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB,
United Kingdom

Laurent Demany

Laboratoire de Neurophysiologie—UMR CNRS 5543, Université Victor Segalen—Bordeaux 2, France

Stephan D. Ewert

Centre for Applied Hearing Research, Acoustic Technology, Ørsted—DTU,
Technical University of Denmark, Denmark

Stanley Sheft

Parnly Hearing Institute, Loyola University of Chicago, 6525 N. Sheridan Road, Chicago, Illinois 60626

Christian Lorenzi

Institut Universitaire de France, Laboratoire de Psychologie Expérimentale—UMR CNRS 8581,
Université René Descartes—Paris 5, France

(Received 13 February 2004; revised 30 December 2004; accepted 30 December 2004)

Recent studies suggest that an auditory nonlinearity converts second-order sinusoidal amplitude modulation (SAM) (i.e., modulation of SAM depth) into a first-order SAM component, which contributes to the perception of second-order SAM. However, conversion may also occur in other ways such as cochlear filtering. The present experiments explored the source of the first-order SAM component by investigating the ability to detect a 5-Hz, first-order SAM probe in the presence of a second-order SAM masker beating at the probe frequency. Detection performance was measured as a function of masker-carrier modulation frequency, phase relationship between the probe and masker modulator, and probe modulation depth. In experiment 1, the carrier was a 5-kHz sinusoid presented either alone or within a notched-noise masker in order to restrict off-frequency listening. In experiment 2, the carrier was a white noise. The data obtained in both carrier conditions are consistent with the existence of a modulation distortion component. However, the phase yielding poorest detection performance varied across experimental conditions between 0° and 180°, confirming that, in addition to nonlinear mechanisms, cochlear filtering and off-frequency listening play a role in second-order SAM perception. The estimated magnitude of the modulation distortion component ranges from 5%–12%. © 2005 Acoustical Society of America.

[DOI: 10.1121/1.1861892]

PACS numbers: 43.66.Ba, 43.66.Dc, 43.66.Mk, 43.66.Nm [NFV]

Pages: 2158–2168

I. INTRODUCTION

The ability of human listeners to detect slow modulations in the amplitude of incoming sounds is generally understood using models based on either a low-pass filter/integrator (Viemeister, 1979; Moore *et al.*, 1988) or a modulation filterbank (MFB) (Dau *et al.*, 1997a, 1997b). In a modulation-masking study, Strickland and Viemeister (1996) first mentioned the ability to perceive the envelope beat produced by the addition of two closely spaced modulation components, which is a slow modulation in envelope *depth*. In parallel, electrophysiological work by Shofner *et al.*

(1996) demonstrated that neural units in the ventral cochlear nucleus of the chinchilla show phase-locked responses both to the frequencies of the modulation primaries and to the modulation difference frequency (i.e., the beat frequency) of two-component sinusoidal amplitude modulation (SAM). These findings appeared to be a challenge for current models of temporal-envelope processing, because there was no energy at the beat frequency in the “physical” modulation spectrum of such stimuli. This was interpreted as evidence for the existence of a nonlinearity along the auditory pathway (e.g., compression, saturation, or thresholding) that generates a *modulation distortion component* at the envelope-beat frequency.

Psychoacoustical evidence consistent with the notion that nonlinear processes are involved in temporal-envelope beat perception was provided by Sheft and Yost (1997) and Moore *et al.* (1999). In both studies, the listeners’ sensitivity to a probe modulator was assessed in the presence of a two-

^{a)}Portions of this work were presented at the 27th Annual Midwinter Research Meeting of the Association for Research in Otolaryngology, Daytona Beach, FL, February 2004.

^{b)}Current address: Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, United Kingdom. Electronic mail: c.fullgrabe@psychol.cam.ac.uk

component masker modulator with modulation difference frequency equal to the probe modulation frequency. The probe and masker modulators were either applied to different sinusoidal carriers (Sheft and Yost, 1997) or to the same sinusoidal or narrowband carrier (Moore *et al.*, 1999). Despite the fact that the physical modulation components of the masker were too remote from the probe to yield modulation-masking effects, the results showed that the presence of the beating masker modulators degraded probe-detection thresholds. Taken together, these studies clearly pointed out the need to incorporate an additional nonlinearity into *each* envelope-processing model to account for the masking data.

The distortion-component hypothesis has received further support from psychoacoustical data on second-order SAM detection (Lorenzi *et al.*, 2001a, 2001b; Füllgrabe and Lorenzi, 2003), that is detection of a sinusoidal modulation applied to the modulation depth of a SAM stimulus. In a masking study, Lorenzi *et al.* (2001b) compared second-order SAM detection for a 5-kHz sinusoidal carrier and a 2-Hz wide narrowband-noise carrier centered at 5 kHz. The inherent random amplitude fluctuations in the narrowband noise would be expected to mask any modulation components in the low modulation frequency region (Dau *et al.*, 1997a). The results revealed that second-order SAM detection thresholds were worse for the narrowband noise than for the sinusoidal carrier only for the lowest second-order modulation frequencies. This finding provided evidence that detection of a distortion component at the envelope-beat frequency contributed to second-order SAM detection.

The potential role of cochlear compression in producing distortion in the internal modulation spectrum has been tested in psychoacoustical studies on second-order SAM detection conducted with hearing-impaired (Tandetnik *et al.*, 2001; Füllgrabe *et al.*, 2003) and cochlear implant listeners (Lorenzi *et al.*, 2004). In both groups, cochlear compression was either severely reduced or completely abolished by cochlear damage. However, second-order temporal modulation transfer functions (TMTFs) relating second-order SAM detection thresholds to second-order modulation frequency were normal or near-normal in shape in these listeners. This demonstrated that cochlear damage has little disruptive effect on the detection of second-order SAM, and, if a distortion component is actually used in second-order SAM detection, it must be generated by a nonlinearity other than cochlear compression.

One way to explore the type of nonlinearity involved in the generation of envelope distortion is to determine the relative phase of the distortion component. In a modulation-masking experiment using a 5-kHz sinusoidal carrier, Verhey *et al.* (2003) systematically varied the phase relationship between the probe modulator and the envelope beat cycle of two- and three-component masker modulators. Irrespective of the masker components' modulation frequencies (ranging from 90 to 360 Hz), maximum masking generally occurred in the antiphase (i.e., 180°) condition, where the maxima in the probe modulator coincided with the minima in the beat cycle of the masker envelope. Moreover, when the probe modulation was in-phase with the envelope beat, masked thresholds were lower than those measured in the unmasked

condition. Taken together, these observations indicated that the generated distortion component was in-phase with the envelope beat cycle. However, with relatively low (i.e., 40 and 45 Hz) modulation components, the phase-dependent masking effects were less pronounced and maximum masking occurred—when observed—at 90° and 180°. In agreement with the hearing-impaired data discussed above, Verhey *et al.* (2003) concluded that the distortion component could not be accounted for by a compressive nonlinearity, which would have yielded a 180° out-of-phase distortion component, that is, maximum masking in the 0° condition. By contrast, the masking data were globally consistent with the notion of “venelope” extraction proposed by Ewert *et al.* (2002). In their functional model of temporal-envelope processing, envelope-beat information is extracted by computing the so-called venelope (the Hilbert envelope of the ac-coupled Hilbert envelope), which is then attenuated and combined with the first-order SAM. In response to complex temporal envelopes, such a mechanism produces an internal modulation component at the envelope-beat frequency that is in-phase with the envelope beat cycle (Ewert *et al.*, 2002; Verhey *et al.*, 2003), and therefore correctly predicts the empirically obtained maximum masking effects for the antiphase condition.

In an attempt to clarify the phase and magnitude of the distortion component, Sek and Moore (2004) measured psychometric functions for detecting a 5-Hz SAM probe applied to a 4-kHz sinusoidal carrier in the presence of a masker modulator with components at 50 and 55 Hz. This time, performance was poorest when the probe modulation was 135° out-of-phase with the masker envelope beat, similar to the phase effect obtained by Verhey *et al.* (2003), but not entirely consistent with predictions based on venelope extraction. In addition, as already apparent in masking data obtained by Verhey *et al.* (2003) using comparable modulation frequencies, the phase effect did vary somewhat across listeners. The effective modulation depth of the distortion component was estimated by Sek and Moore (2004) to be rather weak, at about 3%; however, the authors acknowledged the possibility that the chosen probe phase may not have been optimal to estimate the maximum effective magnitude of the distortion component.

Overall, the preceding studies do not yield fully converging outcomes and showed some degree of inter-listener variability, consistent with the idea that the envelope-beat information may not be converted into a first-order SAM component *via* a single, presumably nonlinear, mechanism. Ewert *et al.* (2002) and Verhey *et al.* (2003) already noted that such a conversion may occur *via* cochlear filtering, provided subjects make use of the output of an auditory filter that is not centered on the carrier frequency. Introduction of a component at the second-order frequency *via* cochlear filtering would mainly occur for sinusoidal carriers. However, it is conceivable that such a conversion also occurs for broadband-noise carriers, especially when the frequency of the first-order SAM carrying the second-order modulation is relatively high (Viemeister, 2003).

Our purpose in the present study was therefore to demonstrate the existence of a “genuine” modulation distortion

component generated by the auditory system in response to second-order masker modulators, and to evaluate the role of other sources of conversion involved in second-order SAM perception. The experiments reported here addressed these issues by investigating the effect of a second-order masker modulator on the detectability of a first-order probe modulator whose modulation frequency corresponded to the periodicity of the masker's envelope beat. Using different carrier signals, the characteristics (i.e., phase and magnitude) of the modulation distortion component were investigated by measuring detection performance as a function of (i) the carrier (or first-order) modulation frequency of the masker; (ii) the phase relationship between the probe and masker modulator; and (iii) the probe modulation depth.

Finally, previous modulation masking studies exploring the phase of the distortion component used procedures with feedback. Using feedback might have led listeners to use *any* cue when performing the detection task (rather than choosing the interval where the modulation depth was greater). This issue seems particularly important when a 2IFC procedure is used in which listeners may change their strategy based on feedback. In their second experiment, Sek and Moore (2004) measured psychometric functions for probe modulation detection as a function of probe depth when no feedback was provided; however, they used a phase relationship between the probe and masker modulator that had been obtained in a previous experiment using feedback. Here, in all masking conditions, no feedback was provided; this was intended to allow us to observe genuine effects of cancellation between the probe modulator and first-order SAM component resulting from conversion, when they were out-of-phase.

II. EXPERIMENT 1: MASKED AND UNMASKED MODULATION DETECTION USING A SINUSOIDAL CARRIER

A. Rationale

Our aim in this experiment was to extend previously published studies conducted with sinusoidal carriers investigating the nature of the nonlinear mechanism involved in the perception of temporal-envelope beat information. Using a 5-kHz sinusoidal carrier, detection performance for a 5-Hz SAM probe was measured in the presence of a second-order SAM masker fluctuating at the probe frequency, as a function of the relative probe phase. Different masker-carrier modulation frequencies were chosen between 32 and 180 Hz to cover the range of modulation frequencies used in earlier studies (Moore *et al.*, 1999; Verhey *et al.*, 2003; Sek and Moore, 2004). Unlike these studies, the sinusoidal carrier was presented within a notched noise to restrict off-frequency listening. To clarify the effects of off-frequency listening, measurements were replicated when the notched-noise masker was removed. The masking study was preceded by a systematic assessment of the listeners' sensitivity to first- and second-order SAM to ensure that the first-order probe and second-order masker modulations were audible in all experimental conditions.

B. First- and second-order TMTFs

1. Listeners

Three listeners aged 20 to 30 years participated in the experiment. All listeners had audiometric thresholds less than 15 dB HL between 0.25 and 8 kHz. One listener was author CF, who had extensive previous experience with psychoacoustic experiments. The other two were volunteers who were paid for their services. Prior to data collection, all listeners received training for about four hours to stabilize thresholds.

2. Stimuli and procedure

All stimuli were generated with a Personal Computer using a 16-bit D/A converter operating at a sampling frequency of 44.1 kHz, and were delivered diotically *via* Sennheiser HD 25-13 earphones. The carrier was a 5-kHz sinusoid presented at 70 dB SPL within a notched-noise masker. The masker was obtained by adding a low-pass noise (cutoff frequency=2917 Hz; roll-off=90 dB/octave) and a high-pass noise (cutoff frequency=7500 Hz; roll-off=90 dB/octave) presented at 20 and 7 dB (rms) below the carrier level, respectively. The corresponding noise spectrum was chosen so that the carrier-to-noise ratio would have been high only for a limited range of auditory filter center frequencies close to the carrier frequency (Alcántara *et al.*, 2003). In all experiments, the stimulus duration was 2 s, including 50-ms rise/fall times shaped using a raised-cosine function. All listeners were tested individually in a sound-attenuating booth.

For the measurement of first-order TMTFs, listeners had to detect SAM applied to the sinusoidal carrier. On each trial, a standard and a target stimulus were presented in random order with an interstimulus interval (ISI) of 1 s. The standard corresponded to the unmodulated carrier. The target corresponded to the carrier that was sinusoidally amplitude modulated at a modulation frequency f_m of 2, 10, 40, 80, or 160 Hz. The target's temporal envelope is given by

$$T_1(t) = 1 + m \sin(2\pi f_m t + \phi), \quad (1)$$

where m is the modulation depth ($0 \leq m \leq 1$), and ϕ is the starting phase of the modulation, randomized in each interval. The overall power was the same in both intervals.

SAM detection thresholds were obtained using an adaptive two-interval, two-alternative forced-choice (2I, 2AFC) procedure with a two-down, one-up stepping rule that estimates the modulation depth m necessary for 70.7% correct detection (Levitt, 1971). Listeners received visual feedback after each trial. The step size was initially 4 dB (in terms of $20 \log m$), and was reduced to 2 dB after the first two reversals. The arithmetic mean of the values of m at the last 10 reversals in a block of 16 reversals was taken as the threshold estimate for that block (reported in %). The final threshold is based on three repeated measures.

For the measurement of second-order TMTFs, listeners had to detect sinusoidal modulation of the modulation depth of first-order SAM. Again, on each trial, a standard and a target stimulus were presented in random order with an ISI of 1 s. The standard had first-order SAM of a given modu-

lation frequency f_m and a fixed modulation depth m of 0.5 (i.e., 50%). The expression describing the standard's temporal envelope was

$$S_2(t) = 1 + 0.5 \sin(2\pi f_m t + \phi), \quad (2)$$

where ϕ represents the starting phase of the modulation, which was randomized for each interval. The target had SAM whose modulation depth was sinusoidally modulated at a (second-order) modulation frequency f'_m . The expression describing the target's temporal envelope was

$$T_2(t) = 1 + [0.5 + m' \sin(2\pi f'_m t + \phi)] \sin(2\pi f_m t + \phi), \quad (3)$$

where m' is the second-order modulation depth ($0 \leq m' \leq 0.5$) and ϕ represents the starting phase of the modulations, randomized for each interval. The overall power was the same in both intervals. The carrier modulation frequency f_m was either 32, 64, or 180 Hz, and the corresponding second-order modulation frequencies f'_m were (i) 2, 5, and 10 Hz, (ii) 2, 5, 10, and 20 Hz, and (iii) 2, 5, 10, 20, 40, and 80 Hz, respectively. Second-order SAM detection thresholds (i.e., m' at threshold) were obtained using an identical 2I, 2AFC psychophysical procedure to that used to determine first-order SAM detection thresholds.

3. Results and discussion

Individual first- (filled squares) and second-order (open symbols) TMTFs for the three listeners are shown in Fig. 1. In agreement with previous experiments using sinusoidal carriers (Viemeister, 1979; Kohlrausch *et al.*, 2000; Lorenzi *et al.*, 2001b; Moore and Glasberg, 2001), first-order SAM detection thresholds show a low- or all-pass characteristic for the range of f_m under study. In all listeners, thresholds are roughly constant ($m \approx 6\%$ or -24 dB expressed as $20 \log m$) at the lowest modulation frequencies; they tend to increase only slightly with f_m for two listeners, but more substantially for listener CF (with a clear increase above 40 Hz). Overall, the thresholds required to detect SAM of a sinusoidal carrier presented in a notched noise are increased by a factor of 3 (10 dB) relative to those obtained in identical but unmasked conditions (Lorenzi *et al.*, 2001b). This difference presumably reflects the contribution of off-frequency listening in the absence of masking noise, which is advantageous because there is less cochlear compression on the high-frequency side of the excitation pattern evoked by the carrier, so the effective magnitude of the modulation is greater on the high-frequency side than at its center (Zwicker, 1956). In addition, some energy from the notched-noise masker may have leaked into the auditory filter centered on the carrier. The fact that the cutoff frequency of the average TMTF is somewhat lower than that reported in previous studies may then be attributed to the increasing masking effect of intrinsic random fluctuations of the notched-noise masker at higher modulation frequencies (Dau *et al.*, 1997a).

The second-order TMTFs show roughly constant thresholds for most second-order modulation frequencies, although a low-pass characteristic with a -3 dB cutoff frequency at or above 20 Hz is apparent when $f_m = 180$ Hz. Moreover, second-order modulation detection thresholds decrease by a

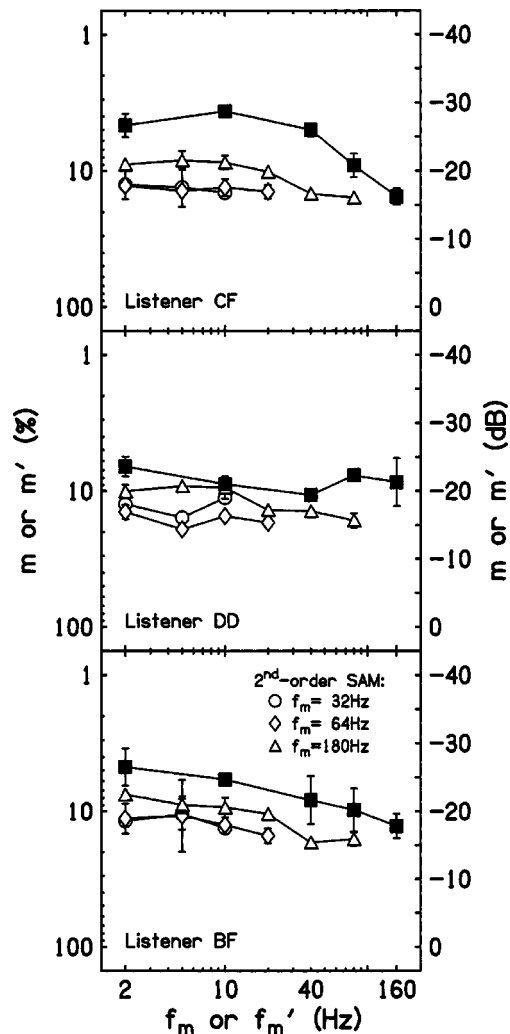


FIG. 1. Individual first- (filled squares) and second-order (open symbols) TMTFs for the three listeners using a 5-kHz sinusoidal carrier presented within a notched-noise masker. First- and second-order modulation depths at threshold m and m' (in %) are plotted as a function of first- and second-order modulation frequency f_m and f'_m , respectively. The right axis shows detection thresholds in dB (in terms of $20 \log m$ or $20 \log m'$). The parameter for the second-order TMTFs is the carrier modulation frequency: $f_m = 32$ (circles), 64 (diamonds), and 180 Hz (triangles). Error bars represent \pm one standard deviation from the mean threshold across three repeated measures.

factor of 1.5 (3 dB) when f_m increases from 32 to 180 Hz. These observations are roughly consistent with data reported by Lorenzi *et al.* (2001b) for similar carrier modulation frequencies using an unmasked 5-kHz carrier. However, as for first-order SAM detection, the average sensitivity to second-order SAM was consistently poorer by approximately a factor of 3 (10 dB) when off-frequency listening was restricted than when off-frequency listening was possible (Lorenzi *et al.*, 2001b).

C. Detection of a 5-Hz probe modulator in the presence of a 5-Hz, second-order SAM masker: Phase effects

1. Stimuli and procedure

The three listeners used to assess SAM-detection thresholds participated in this experiment. The detectability of a

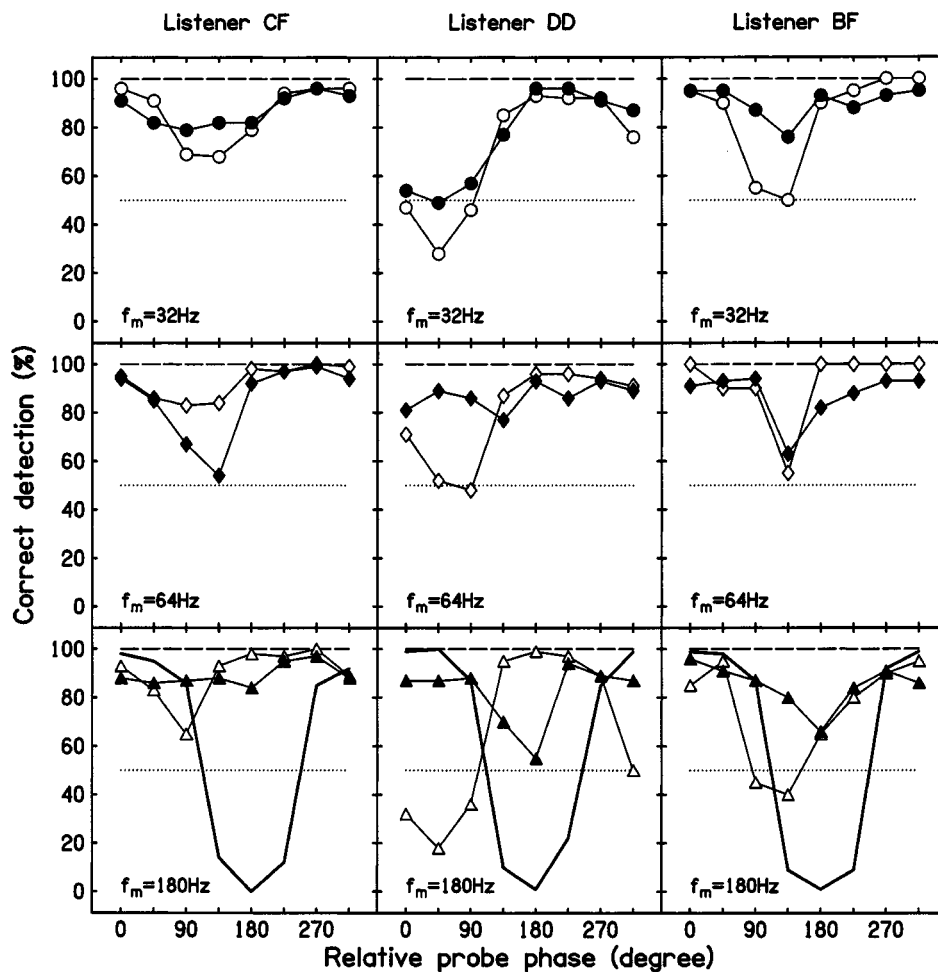


FIG. 2. Individual detection scores for 20%, 5-Hz probe modulation of a 5-kHz sinusoidal carrier presented within a notched-noise masker, measured in three listeners (columns). Performance is plotted as a function of the phase relationship between the probe and (i) a 5-Hz, second-order SAM masker for carrier modulation frequencies of 32 (open circles, upper row), 64 (open diamonds, middle row), and 180 Hz (open triangles, lower row); and (ii) a 20%, 5-Hz SAM masker (bold line, lower row). In each panel, unmasked probe detection scores are indicated by the long-dashed line. Filled symbols correspond to masked detection performance when the notched noise was removed. The dotted line represents performance at chance level.

20%, 5-Hz SAM probe was measured using a 2I, 2AFC procedure with constant stimuli. The probe was applied to a 5-kHz sinusoidal carrier embedded within the notched-noise masker, and presented either:

- (1) Alone, without any additional SAM masker (unmasked detection condition).
- (2) In the presence of a second-order SAM masker that was applied to the sinusoidal carrier in both intervals. The first-order modulation frequency of the masker was 32, 64, or 180 Hz, while the second-order modulation frequency was fixed at 5 Hz. The first- and second-order modulation depths of the masker modulator were set to 40%. All masker modulation components started in sine phase. Eight phase relationships between the probe and masker modulators were used: 0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315°. In this condition, performance was also assessed without the notched-noise masker.
- (3) In the presence of a first-order SAM masker applied to the sinusoidal carrier in both intervals. The probe and masker modulators had identical modulation depths of 20% and modulation frequencies of 5 Hz. As with the second-order masker modulator, the phase relationship between probe and masker modulators was shifted in steps of 45° from 0° to 315°.

The inspection of first- and second-order modulation detection thresholds reported above indicates that in the present

experimental conditions, all first- and second-order modulation components were audible when presented individually.

Prior to data collection, listeners received 20 practice trials in each experimental condition. Detection scores presented here are based on 100 trials.

2. Results and discussion

Figure 2 shows individual probe detection scores for the three listeners (columns) as a function of the phase relationship between the probe and the masker modulator. Using a notched-noise masker, detection performance is perfect for all listeners when the 20% probe modulator is presented alone (long-dashed line), and decreases when a second-order masker modulator (open symbols) is added, despite the fact that the individual modulation components of the latter are remote in frequency from the probe modulator. This can be explained in terms of an interaction of the probe modulator with a modulation distortion component produced in response to the second-order modulator: sometimes the probe modulator would have been in-phase with the distortion component, so that they added, and sometimes the two would have been out-of-phase, so that they cancelled. Since no feedback was provided in the masked conditions, a drop of detection performance below chance level for a given phase relationship reflects the fact that listeners tended to choose the standard instead of the target interval as the one

containing the more salient 5-Hz SAM, consistent with the notion that the probe modulation was (partially) cancelled.

The results show that the phase relationship between the probe and masker modulators leading to the poorest performance does not vary with masker-carrier modulation frequency *within* the same listener. However, as already observed to a lesser degree in previous studies (Verhey *et al.*, 2003; Sek and Moore, 2004), phase effects do vary *across* listeners with maximum masking occurring for probe phases of 45° to 135°. Such a shift is not expected if there were a *single* static nonlinearity, as postulated in the literature (Moore *et al.*, 1999; Verhey *et al.*, 2003), but may be explained by the existence of several static nonlinearities introducing distortion into the internal representation of complex temporal envelopes. The exact contribution of each nonlinearity would have to be listener-specific to account for the observed inter-listener variability.

The replication of the experiment without the notched-noise masker (filled symbols) yields a similar pattern of results for $f_m = 32$ Hz in all listeners, but shows important differences for $f_m = 64$ and 180 Hz for two out of the three listeners. This suggests that, for the lowest f_m under study, the phase effect is not substantially influenced by off-frequency listening. In contrast, but in agreement with recent data reported by Sek and Moore (2004), the maximum decrease in performance for $f_m = 64$ Hz now occurs consistently around 135°. For listener DD, the phase at which performance is worst is shifted by +45°, relative to what was obtained in the notched-noise condition. An even stronger discrepancy between the results obtained with and without the notched noise is observed for $f_m = 180$ Hz. Here, two listeners show a clear shift of the phase yielding the poorest performance when the notched-noise masker is removed. The worst detection scores occur at 180° in listeners DD and BF, while no phase-dependent masking effect is observed anymore in listener CF. In comparison, Verhey *et al.* (2003) reported the poorest detection performance systematically occurring at 180° for a wide range of f_m , except for the lowest f_m under study.

The fact that probe detectability is not only dependent on phase but also varies with f_m when the notched noise is removed suggests that listeners use off-frequency listening in certain conditions. Off-frequency listening may influence performance in two ways. First, although there is no modulation component at the second-order frequency in the modulation spectrum of the stimulus, the envelope at the output of an auditory filter tuned away from the signal frequency may contain such a component (e.g., Ewert *et al.*, 2002). Second, basilar-membrane compression is strong for auditory filters tuned close to the signal frequency, but weak for the more remote filters. This might affect the balance between different sources of nonlinearity. This interpretation seems however less likely given the observation that hearing-impaired and cochlear implant listeners show quasi-normal sensitivity to second-order SAM (Tandetnik *et al.*, 2001; Füllgrabe *et al.*, 2003; Lorenzi *et al.*, 2004).

In the present experiment, the clearest indication of conversion *via* cochlear filtering and off-frequency listening is obtained for two out of the three listeners at $f_m = 180$ Hz.

Here, the phase dependence of probe detectability for a 5-Hz, first-order SAM masker (bold line, lower row) is nearly identical to that measured for the second-order SAM masker.

The present results obtained without a notched-noise masker are at least partially consistent with phase effects previously reported by Verhey *et al.* (2003) and Sek and Moore (2004), suggesting that off-frequency listening might also have occurred in these studies. This assumption is emphasized by the observation that Verhey *et al.* (2003) failed to find clear phase-dependent masking effects when a masker modulator with low (i.e., 40 and 45 Hz) components was used. Thus, the masking effects reported by Verhey *et al.* (2003) and Sek and Moore (2004) may not *solely* reflect an interaction between the probe and a genuine modulation distortion component.

It is noteworthy that, when a carrier modulation frequency of 64 or 180 Hz and no notched-noise masker was used, listeners reported the presence of a weak high-pitch tone in the standard interval only. Therefore, despite the fact that listeners received instructions to choose the interval containing the most salient 5-Hz modulation, listeners may have used both cues to perform the detection task when no notched noise was used.

III. EXPERIMENT 2: MASKED AND UNMASKED MODULATION DETECTION USING NOISE CARRIERS

A. Rationale

The preceding experiment showed that nonlinear mechanisms are indeed involved in the conversion of second-order SAM into a first-order SAM component. However, the results also suggested that, under certain conditions, second-order SAM is converted into a first-order SAM component *via* cochlear filtering. A similar conversion may occur for broadband-noise carriers, especially when the carrier modulation frequency is high (i.e., greater than the bandwidth of the excited auditory filters). Simulations of the output of cochlear filters were therefore performed to evaluate the existence of this source of conversion for broadband-noise stimuli.

Figure 3 shows the responses of an array of gammatone filters (Patterson *et al.*, 1987) to a 5-Hz, second-order SAM noise (0–20 kHz). Two carrier modulation frequencies [$f_m = 64$ (diamonds) and 180 Hz (triangles)] falling within the frequency range of modulation primaries used in previous psychoacoustical studies (Moore *et al.*, 1999; Verhey *et al.*, 2003; Sek and Moore, 2004), and a high carrier modulation frequency of 2 kHz (circles) were used. Modulation depths m and m' were set to 40%. Also shown are responses to an unmodulated noise (filled squares) and a 8%, 5-Hz (first-order) SAM noise (filled inverted triangles). The magnitude (left panel) and phase (right panel) of the 5-Hz (first-order) SAM component observed at the filter outputs are shown as a function of filter center frequency (CF). For all three carrier modulation frequencies, the data confirm that the responses of cochlear filters (modeled here as linear filters) contain indeed a first-order modulation component at 5 Hz. When $f_m = 2$ kHz, a 5-Hz modulation component is observed at all CFs and its magnitude and phase are about the same as those

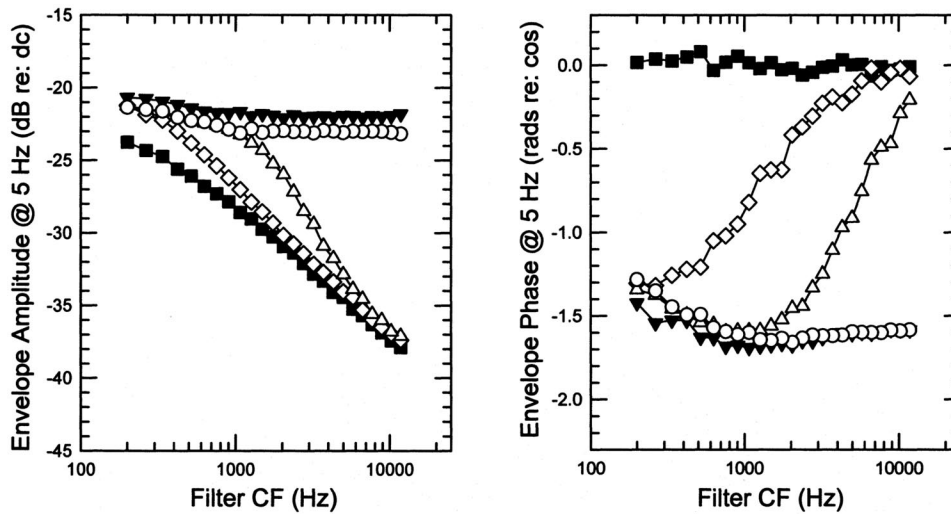


FIG. 3. Simulations of outputs of linear auditory filters in response to second-order SAM noise (0–20 kHz) with carrier modulation frequencies of 64 (diamonds), 180 (triangles), and 2000 Hz (circles). Also are shown responses to an unmodulated (filled squares) and a 8%, 5-Hz SAM noise (filled inverted triangles). The magnitude (left panel) and phase (right panel) of the 5-Hz (first-order) SAM component at the outputs of auditory filters are shown as a function of the filter center frequency (CF). Each data point corresponds to the mean of 2500 presentations of each stimulus. Note that, in referencing the measured phase at each frequency to a cosine at that frequency, a sine function would have a phase value of -1.57 ($-\pi/2$) rad re:cos.

obtained with 8%, 5-Hz (first-order) SAM. In other words, the maxima of this first-order modulation component coincide with the maxima in the envelope beat cycle of the second-order SAM. For the lower carrier modulation frequencies, the 5-Hz modulation component only occurs at the outputs of filters with low CFs (i.e., with ERBs smaller than roughly $2 \times f_m$) and its magnitude increases progressively as CF decreases. The 5-Hz modulation component at the outputs of auditory filters with CF-below approximately 300 Hz (when $f_m = 64$ Hz) and 1 kHz (when $f_m = 180$ Hz) is roughly in sine phase (i.e., in-phase with the second-order SAM). For filters with higher CFs, the modulation component shifts progressively out-of-phase with the second-order modulator.

The following experiments were designed to investigate the relative contribution of each source of conversion (i.e., cochlear filtering and auditory nonlinearities) in the case of broadband-noise carriers. Again, probe detectability was measured behaviorally as a function of the phase relationship between the probe modulator and a second-order SAM masker fluctuating at the probe frequency. The masking data for a carrier modulation frequency of 64 Hz were compared using a white noise and 6-kHz wide bandlimited noise (4–10 kHz). This comparison allowed determining whether cancellation effects could still be observed for broadband-noise carriers, even though conversion through cochlear filtering did not occur in this case (cf. the left panel of Fig. 3). This experiment also allowed assessing the relative phase of the genuine modulation distortion component when using a noise carrier. Additional masking data were collected with the white-noise carrier for the other two carrier modulation frequencies used in the simulations (180 Hz and 2 kHz) for which conversion through cochlear filtering was shown to occur in most or all filters. Finally, probe detectability was assessed as a function of probe modulation depth in the white-noise carrier condition in order to quantify the magnitude of the generated first-order SAM component for each masker-carrier modulation frequency. As in the first experi-

ment, the modulation-masking experiments were preceded by a systematic assessment of the listeners' sensitivity to first- and second-order SAM white noise.

B. First- and second-order TMTFs

1. Listeners

Four listeners ranging in age between 18 and 30 years were tested. One of them was author CF and the other three were paid volunteers. All listeners had absolute thresholds less than 15 dB HL at all audiometric frequencies. Practice of 3 h was given prior to data collection.

2. Stimuli and procedure

The apparatus, procedure, and stimulus parameters were the same as for experiment 1, except that (i) a different set of earphones (Sennheiser HD 565) was used; (ii) SAM was applied to a white-noise carrier; (iii) first-order SAM detection thresholds were obtained for $f_m = 2, 10, 40, 160,$ and 2000 Hz; and (iv) second-order SAM detection thresholds were obtained for carrier modulation frequencies f_m of 64, 180, and 2000 Hz; the second-order modulation frequencies f'_m were 2, 5, 10, and 20 Hz when $f_m = 64$ Hz, 2, 5, 10, 20, 40, and 80 Hz when $f_m = 180$ Hz, and 2, 5, 10, 20, 40, 80, and 160 Hz when $f_m = 2000$ Hz.

3. Results and discussion

Figure 4 shows individual first- (filled squares) and second-order (open symbols) TMTFs for the four listeners. Consistent with previous data obtained with broadband-noise carriers (Viemeister, 1979; Bacon and Viemeister, 1985; Lorenzi *et al.*, 2001a), the first-order TMTFs are low-pass in shape with thresholds of about 3% (-30 dB) for the lowest values of f_m , and a -3 dB cutoff frequency below 40 Hz. Note that listeners required modulation depths ranging from 45% to 51% (-6.9 dB to -5.8 dB) to reach the threshold

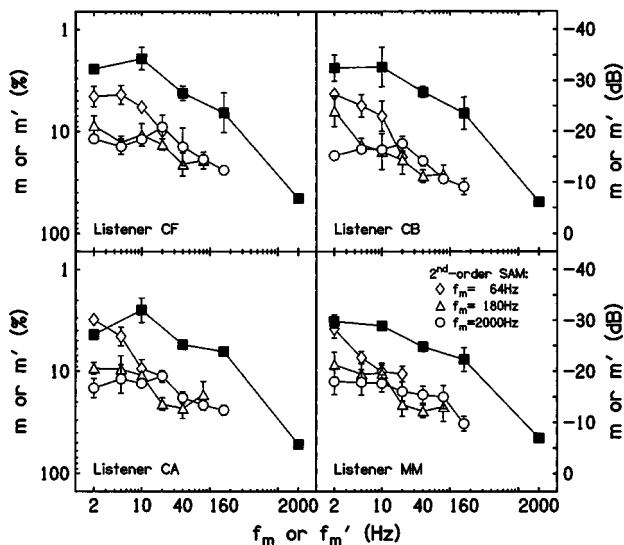


FIG. 4. Individual first- (filled squares) and second-order (open symbols) TMTFs for the four listeners using a white-noise carrier. The carrier modulation frequency of the second-order SAM was 64 (diamonds), 180 (triangles), and 2000 Hz (circles). Otherwise as in Fig. 1.

criterion of 70.7% correct for 2-kHz SAM. This means that, in the case of second-order SAM with a 50% carrier SAM (as was used to measure second-order TMTFs), the first-order modulation was approximately at threshold.

Similar to data reported by Lorenzi *et al.* (2001a), second-order TMTFs with $f_m = 64$ (diamonds) and 180 Hz (triangles) show a low-pass characteristic, with average sensitivity decreasing and cutoff frequency increasing with carrier modulation frequency. Second-order TMTFs obtained with $f_m = 2$ kHz (circles) show constant thresholds of about 13% (-18 dB) at the lowest second-order SAM frequencies, and increase progressively beyond $f'_m = 20$ Hz. The shapes of these second-order TMTFs mirror quite closely the first-order TMTFs, suggesting that the second-order SAM was detected using the same (but attenuated) cue as used for first-order SAM detection.

C. Detection of a 5-Hz probe modulator in the presence of a 5-Hz, second-order SAM masker: Phase effects

1. Stimuli and procedure

For the same four listeners, the detectability of a 5-Hz SAM probe applied to a white-noise carrier was measured, following the same protocol as in experiment 1, except that (i) Sennheiser HD 565 earphones were used; (ii) the probe modulation depth was fixed at a supra-threshold level of 5%, and therefore the modulation depth of the first-order SAM masker was also 5%; and (iii) the second-order masker-carrier modulation frequency was either 64, 180, or 2000 Hz. As for the experiment conducted with the sinusoidal carrier, the detection data shown in Fig. 4 indicate that all first- and second-order modulation components were audible when presented individually.

Two (CF,CB) out of the four listeners were tested when $f_m = 64$ and 2000 Hz using a 6-kHz wide broadband noise as a carrier instead of an unfiltered white noise. The

bandlimited-noise carrier was obtained by bandpass filtering a white noise between 4 and 10 kHz (roll-off=120 dB/octave). The spectrum level was fixed at 32 dB. Complementary low and high unmodulated noise flankers (with frequency ranges of 0–4 and 10–22.5 kHz) were added to the modulated bandlimited noise with spectrum levels of 34 and 29 dB, respectively. This was done to prevent listeners from performing the detection task using spectral cues or combination tones (Wiegrebe and Patterson, 1999).

2. Results and discussion

Figure 5 shows individual detection scores for the white-noise (open symbols) and bandlimited-noise (filled symbols) conditions. Data for the three carrier modulation frequencies (rows) are shown as a function of relative probe phase.

As expected based on the SAM-detection data (Fig. 4), near-perfect detection performance is observed in the white-noise condition when the 5% probe modulator is presented alone (long-dashed line). Detection performance drops to about 73% correct when the carrier is bandlimited (short-dashed line). Overall, adding a second-order masker modulator to the SAM probe degrades performance for both types of noise carriers. Moreover, as in the previous masking experiment using a sinusoidal carrier, probe detection varies with the phase relationship between the probe and masker modulators, but the probe phases giving maximum and minimum detectability depend on the masker-carrier modulation frequency. In contrast to the previously collected data, all phase effects are highly consistent across listeners.

The upper row shows the data for a second-order SAM masker with $f_m = 64$ Hz in the white-noise and bandlimited-noise conditions. In both cases, the detectability of the probe varies in essentially the same way as a function of relative probe phase. Minimum and maximum detectability occur for $0/45^\circ$ and $180/225^\circ$, respectively. The fact that the overall pattern of results is very similar for both noise carriers can be taken as evidence that cochlear filtering does not contribute to the observed phase effects when $f_m = 64$ Hz. Consequently, the observed cancellation effects demonstrate the existence of a genuine distortion component with a phase of $180/225^\circ$. Surprisingly, this phase is neither consistent with previously published data (in which off-frequency may have occurred) nor with the results from experiment 1 (in which off-frequency listening was, however, precluded). This discrepancy concerning the phase of the modulation distortion component between studies using noise and sinusoidal carriers remains unexplained.

The middle row shows the masking data for the white-noise condition when f_m is 180 Hz. Maximum and minimum scores occur on average at about 270° and $90/135^\circ$, respectively. However, in this case, performance does not fall below chance (except for listener CA). Another pattern is obtained for $f_m = 2$ kHz (open circles, lower row), where maximum scores are observed for phase relationships of 0° , 45° , and 315° , and performance consistently drops to a minimum at 180° . The fact that performance for the out-of-phase condition approaches closely the worst possible detection score suggests that the probe modulator and the first-order SAM component resulting from conversion are of similar

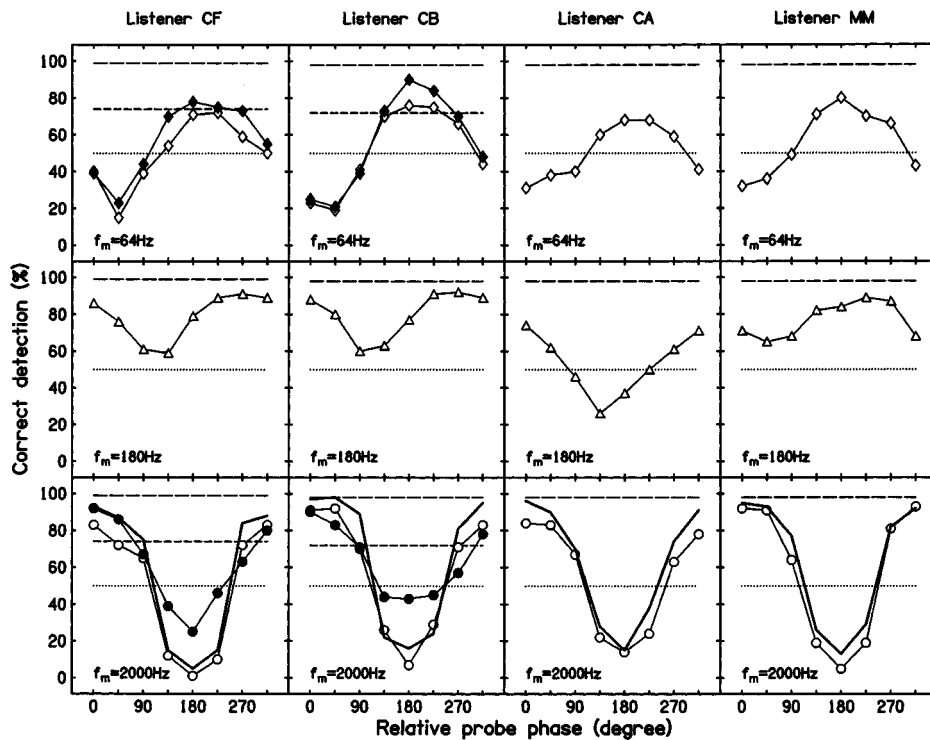


FIG. 5. Individual detection scores for 5%, 5-Hz probe modulation of a white-noise (open symbols) and bandlimited-noise (filled symbols) carrier, measured for four listeners (columns). Carrier modulation frequency of the second-order SAM masker was fixed at 64 (diamonds, upper row), 180 (triangles, middle row), or 2000 Hz (circles, lower row). The modulation depth of the 5-Hz SAM masker (bold line, lower row) was set to 5%. Unmasked probe detection scores are indicated for the white-noise (long-dashed line) and bandlimited-noise (short-dashed line) conditions. The dotted line represents performance at chance level.

magnitude and approximately in antiphase, so that they almost cancel each other out. In agreement with this idea, the pattern of results for the 5%, first-order SAM masker (bold line, lower row), is nearly identical to that for the second-order SAM white noise with $f_m = 2$ kHz. These data appear consistent with what would be predicted if detection performance were *mainly* determined by the interaction between the probe modulator and the 8%, first-order SAM component appearing at the output of cochlear filters at all CFs (cf. Fig. 3). The progressive change in the phase yielding maximum cancellation with an increase in f_m from 64 to 2 kHz is therefore in line with the notion that the contribution of the first-order SAM component resulting from cochlear conversion increases with carrier modulation frequency.¹

D. Detection of a 5-Hz probe modulator in the presence of a 5-Hz, second-order SAM masker: Effects of probe modulation depth

1. Stimuli and procedure

Psychometric functions were measured for detecting a 5-Hz SAM probe in the presence of a second-order SAM masker as a function of the probe modulation depth in the white-noise condition. Three of the original four listeners were tested. The apparatus, procedure, and stimuli were identical to those described in Sec. III C, except that, for each masker-carrier modulation frequency, the probe phase was fixed for each listener at the value yielding the poorest performance in the previous experiment (cf. Fig. 5). We assumed that, for this value, the probe modulator and the first-order SAM component resulting from conversion of envelope-beat information were 180° out-of-phase. The modulation depth of the probe modulator was either 1.25%, 2.5%, 5%, 10%, or 20%.

2. Results and discussion

Figure 6 shows individual detection scores for the three listeners as a function of probe modulation depth for each carrier modulation frequency of the second-order masker. The psychometric functions obtained with $f_m = 64$ Hz (diamonds) and 2000 Hz (circles) are nonmonotonic and show a u-shape, with worst performance occurring for intermediate probe depths of 5% and 10%. When $f_m = 180$ Hz (triangles), the deleterious effect of the masker is generally reduced; the psychometric functions increase monotonically with increasing probe depth with worst performance occurring consistently for probe depths of 1.25% to 5%. Taken together, these patterns of results suggest that the magnitude of the first-order conversion component is strictly below 20%. This upper limit of the estimated magnitude is somewhat higher than the 3% value reported in earlier studies using a sinusoidal audio carrier (Moore *et al.*, 1999; Sek and Moore, 2004).

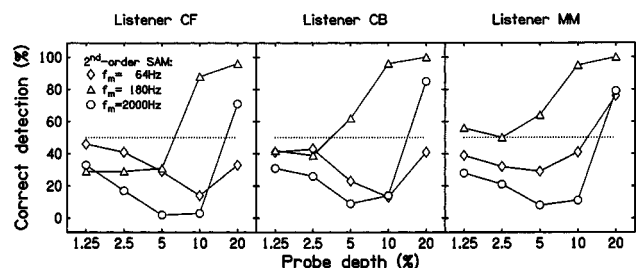


FIG. 6. Individual masked detection scores for 5%, 5-Hz probe modulation of a white-noise carrier as a function of probe modulation depth, measured for three listeners. The masker was 5-Hz, second-order SAM with carrier modulation frequency of 64 (diamonds), 180 (triangles), and 2000 Hz (circles). The phase relationship between the probe and each second-order SAM masker was fixed in such a way as to correspond to the phase condition yielding the maximum degradation of performance in Fig. 5. Performance at chance level is indicated by the dotted line.

TABLE I. Estimated magnitude m_c in % (or in terms of $20 \log m_c$, given in parentheses) of the first-order SAM component, yielding the highest correlation coefficient r between predicted and observed probe detection performance. The masker-carrier modulation frequency f_m was 64, 180, or 2000 Hz.

	$f_m = 64$ Hz		$f_m = 180$ Hz		$f_m = 2000$ Hz	
	m_c	r	m_c	r	m_c	r
CF	12 (-18.4)	0.99	2 (-34)	0.94	8 (-21.9)	0.97
CB	11 (-19.2)	0.97	2 (-34)	0.95	7 (-23.1)	0.99
MM	5 (-26)	0.99	2 (-34)	0.96	7 (-23.1)	0.99

E. Simulations

Taken together, the empirical results and simulations appear consistent with the notion that for broadband-noise carriers: (i) both cochlear filtering and nonlinear mechanisms can introduce first-order SAM components into the internal representation of the envelope of a second-order SAM stimulus; (ii) these first-order SAM components—when present simultaneously—interact according to their own phase and magnitude, resulting in a “compound” first-order SAM component; and (iii) detectability of a first-order SAM probe in the presence of a second-order SAM masker is determined by the specific interaction of the (compound or not) first-order SAM component and the first-order SAM probe. This “cancellation hypothesis” was further tested using an approach similar to that taken by Sek and Moore (2004) to model their detection data.

This model assumes that listeners perform the probe detection task on the basis of a change in the depth of the 5-Hz modulation. In other words, listeners distinguish the 5-Hz SAM conversion component of depth m_c in the standard interval from the 5-Hz SAM in the target interval of magnitude m_{sum} , corresponding to the vector sum of the first-order SAM component (of depth m_c) and the probe modulation (of depth m_p). For some probe phases, the first-order SAM component and the probe modulation tend to cancel, leading to a small value of m_{sum} . Since performance is assumed to be monotonically related to $m_{\text{sum}} - m_c$, poor performance is expected in this condition.

To test this prediction, for each listener a starting value was used for m_c . Assuming that the first-order SAM component was 180° out-of-phase with the probe modulation in the last experiment (for each listener, the probe phase was fixed at the value yielding the poorest performance), the value of m_{sum} was calculated for each probe modulation depth used to obtain the psychometric functions in Sec. III D. The correlation of the d' values (computed from the obtained percent correct scores) with the values of $m_{\text{sum}} - m_c$ was then determined, and the value of m_c was systematically varied to determine the value giving the highest correlation. The resulting values of m_c and correlation coefficients are given in Table I. Scatter plots of the values of d' against the values of $m_{\text{sum}} - m_c$ were fitted with regression lines, and those lines were used to generate predicted values of d' for each probe modulation depth (data not shown). There was no

evidence of any systematic discrepancy between the predicted and observed d' values.

As shown in Table I, the values of m_c ranged from 2% to 12% (i.e., -34 dB to -18.4 dB expressed as $20 \log m_c$), with correlations between d' and $m_{\text{sum}} - m_c$ ranging from 0.94 to 0.99). This suggests that, when $f_m = 64$ or 2000 Hz, the first-order SAM component is easily detectable (cf. first-order SAM detection thresholds shown in Fig. 4). In contrast, when $f_m = 180$ Hz, the first-order SAM component is barely audible, consistent with earlier estimates of the effective magnitude of the modulation distortion component using a sinusoidal carrier (Moore *et al.*, 1999; Sek and Moore, 2004). Combining these magnitude estimates and the simulated outputs of cochlear filters suggests that the magnitude of the genuine modulation distortion component (i.e., for low carrier modulation frequencies where cochlear filtering does not play a role) ranges from 5% to 12%. In addition, in the case of the 2 kHz carrier modulation frequency, the first-order SAM component seems to result entirely from cochlear conversion since its estimated magnitude corresponds precisely to the magnitude of the first-order SAM component observed at the output of cochlear filters in response to a second-order SAM with $f_m = 2000$ Hz.

IV. CONCLUSIONS

The following conclusions can be drawn from the reported data.

- (1) Consistent with earlier modulation masking studies, the detectability of a SAM probe was degraded in the presence of second-order SAM masker beating at the probe frequency, even though all components in the modulation spectrum of the masker were remote from the probe frequency. The effect depended on the stimulus parameters, such as masker-carrier modulation frequency, phase relationship between the probe modulator and the second-order SAM masker, and probe modulation depth, as well as the type of carrier signal.
- (2) When a sinusoidal carrier is used, the results indicate that off-frequency listening influences detection performance, revealing that, for higher carrier modulation frequencies, cochlear filtering is potentially involved in the conversion of the beat of complex envelopes into a first-order modulation component. However, the fact that masking effects were still observed when off-frequency listening was precluded argues for the existence of a genuine modulation distortion component. In this condition, the phase effects varied across listeners from 45° to 135° , suggesting that several nonlinear mechanisms may be involved in the generation of modulation distortion.
- (3) When a white-noise carrier is used, combined empirical results and computer simulations of the outputs of auditory filters indicate, again, that cochlear filtering is potentially involved in the conversion of the beat of complex envelopes into a first-order modulation component, but mainly for high carrier modulation frequencies (e.g., 2 kHz). The masking effects observed at lower carrier modulation frequencies when no other sources of conversion play a role (e.g., cochlear filtering or sideband

reflection) are in line with the existence of a modulation distortion component with a phase of $180/225^\circ$ relative to the envelope beat.

- (4) When a white-noise carrier is used, the magnitude of the first-order modulation presumably resulting from one or the interaction of several sources of conversion ranges from 1% to 12% when first- and second-order modulation depths are fixed at 40%. This suggests that this component may be largely audible under certain stimulus conditions. The magnitude of the genuine modulation distortion component is estimated between 5% and 12% when the carrier modulation frequency is 64 Hz.

ACKNOWLEDGMENTS

This research was supported by a MENRT grant to C. Füllgrabe, and a grant from the Institut Universitaire de France to C. Lorenzi. The work of B. C. J. Moore was supported by the MRC (UK). The work of S. Sheft was supported by the NIH. We thank the editor Neal F. Viemeister and two anonymous reviewers for constructive comments on earlier versions of this manuscript.

¹In addition to cochlear filtering, conversion may also occur physically for broadband-noise carriers through intermodulation introduced by sideband reflection. Modulation adds upper and lower sidebands to each spectral component of the carrier with a carrier-to-sideband frequency separation equal to the frequency of the specific modulator component. When sideband placement is either below dc or above the Nyquist frequency of a digital signal, the sidebands mirror about those respective points. A consequence of this sideband reflection is intermodulation in the modulation domain. That is, components are added to the modulation spectrum at the sum and the difference terms between integer multiples of the component frequencies of the modulator. In the case of second-order SAM, this includes a significant component at the second-order modulation frequency. With a white-noise carrier and a relatively high frequency of the first-order SAM carrying the second-order modulation, noticeable sideband reflection can occur. However, since the overall patterns of results in the bandlimited-noise condition, in which reflected sidebands were removed and/or masked (Fig. 5, filled symbols), are similar to those obtained in the white-noise condition (Fig. 5, open symbols), it may be concluded that sideband reflection did not influence the present results obtained with the white-noise carrier.

Alcántara, J. I., Moore, B. C. J., Glasberg, B. R., and Wilkinson, A. J. K. (2003). "Phase effects in masking: Within-Versus across-channel processes," *J. Acoust. Soc. Am.* **114**, 2158–2166.

Bacon, S. P., and Viemeister, N. F. (1985). "Temporal modulation transfer functions in normal-hearing and hearing-impaired subjects," *Audiology* **24**, 117–134.

Dau, T., Kollmeier, B., and Kohlrausch, A. (1997a). "Modeling auditory processing of amplitude modulation: I. Modulation detection and masking with narrow-band carriers," *J. Acoust. Soc. Am.* **102**, 2892–2905.

Dau, T., Kollmeier, B., and Kohlrausch, A. (1997b). "Modeling auditory processing of amplitude modulation: II. Spectral and temporal integration in modulation detection," *J. Acoust. Soc. Am.* **102**, 2906–2919.

Ewert, S. D., Verhey, J. L., and Dau, T. (2002). "Spectro-temporal processing in the envelope-frequency domain," *J. Acoust. Soc. Am.* **112**, 2921–2931.

Füllgrabe, C., and Lorenzi, C. (2003). "The role of envelope beat cues in the detection and discrimination of second-order amplitude modulation," *J. Acoust. Soc. Am.* **113**, 49–52.

Füllgrabe, C., Meyer, B., and Lorenzi, C. (2003). "Effect of cochlear damage on the detection of complex temporal envelopes," *Hear. Res.* **178**, 35–43.

Kohlrausch, A., Fassel, R., and Dau, T. (2000). "The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers," *J. Acoust. Soc. Am.* **108**, 723–734.

Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.

Lorenzi, C., Soares, C., and Vönnert, T. (2001a). "Second-order temporal modulation transfer functions," *J. Acoust. Soc. Am.* **110**, 1030–1038.

Lorenzi, C., Simpson, M. I. G., Millman, R. E., Griffiths, T. D., Woods, W. P., Rees, A., and Green, G. G. R. (2001b). "Second-order modulation detection thresholds for pure-tone and narrow-band noise carriers," *J. Acoust. Soc. Am.* **110**, 2470–2478.

Lorenzi, C., Sibellas, J., Füllgrabe, C., Gallégo, S., Fugain, C., and Meyer, B. (2004). "Effects of amplitude compression on first- and second-order modulation detection thresholds in cochlear implant listeners," *Int. J. Audiol.* **43**, 264–270.

Moore, B. C. J., and Glasberg, B. R. (2001). "Temporal modulation transfer functions obtained using sinusoidal carriers with normally hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **110**, 1067–1073.

Moore, B. C. J., Sek, A., and Glasberg, B. R. (1999). "Modulation masking produced by beating modulators," *J. Acoust. Soc. Am.* **106**, 908–918.

Moore, B. C. J., Glasberg, B. R., Plack, C. J., and Biswas, A. K. (1988). "The shape of the ear's temporal window," *J. Acoust. Soc. Am.* **83**, 1102–1116.

Patterson, R. D., Nimmo-Smith, J., Holdsworth, J., and Rice, P. (1987). "An efficient auditory filterbank based on the gammatone function," paper presented at a *Meeting of the IOC Speech Group on Auditory Modelling at RSRE*.

Sek, A., and Moore, B. C. J. (2004). "Estimation of the level and phase of the simple distortion tone in the modulation domain," *J. Acoust. Soc. Am.* **116**, 3031–3037.

Sheft, S., and Yost, W. A. (1997). "Modulation detection interference with two-component masker modulators," *J. Acoust. Soc. Am.* **102**, 1106–1112.

Shofner, W. P., Sheft, S., and Guzman, S. J. (1996). "Responses of ventral cochlear nucleus units in the chinchilla to amplitude modulation by low-frequency, two-tone complexes," *J. Acoust. Soc. Am.* **99**, 3592–3605.

Strickland, E. A., and Viemeister, N. F. (1996). "Cues for discrimination of envelopes," *J. Acoust. Soc. Am.* **99**, 3638–3646.

Tandetnik, S., Garnier, S., and Lorenzi, C. (2001). "Measurement of first- and second-order modulation detection thresholds in listeners with cochlear hearing loss," *Br. J. Audiol.* **35**, 355–364.

Verhey, J. L., Ewert, S. D., and Dau, T. (2003). "Modulation masking produced by complex tone modulators," *J. Acoust. Soc. Am.* **114**, 2135–2146.

Viemeister, N. F. (1979). "Temporal modulation transfer functions based upon modulation thresholds," *J. Acoust. Soc. Am.* **66**, 1364–1380.

Viemeister, N. F. (2003) (personal communication).

Wiegand, L., and Patterson, R. D. (1999). "Quantifying the distortion products generated by amplitude-modulated noise," *J. Acoust. Soc. Am.* **106**, 2709–2718.

Zwicker, E. (1956). "Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs (The foundations for determining the information capacity of the auditory system)," *Acustica* **6**, 356–381.

The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners^{a)}

Tanya L. Arbogast,^{b)} Christine R. Mason, and Gerald Kidd, Jr.
*Hearing Research Center and Communication Disorders, Boston University, 635 Commonwealth Avenue,
Boston, Massachusetts 02215*

(Received 20 April 2004; revised 30 September 2004; accepted 5 January 2005)

The ability to understand speech in a multi-source environment containing informational masking may depend on the perceptual arrangement of signal and masker objects in space. In normal-hearing listeners, Arbogast *et al.* [J. Acoust. Soc. Am. **112**, 2086–2098 (2002)] found an 18-dB spatial release from a primarily informational masker, compared to 7 dB for a primarily energetic masker. This article extends the earlier work to include the study of listeners with sensorineural hearing loss. Listeners performed closed-set speech recognition in two spatial conditions: 0° and 90° separation between signal and masker. Three maskers were tested: (1) the different-band sentence masker was designed to be primarily informational; (2) the different-band noise masker was a control for the different-band sentence; and (3) the same-band noise masker was designed to be primarily energetic. The spatial release from the different-band sentence was larger than for the other maskers, but was smaller (10 dB) for the hearing-impaired group than for the normal-hearing group (15 dB). The smaller benefit for the hearing-impaired listeners can be partially explained by masker sensation level. However, the results suggest that hearing-impaired listeners can use the perceptual effect of spatial separation to improve speech recognition in the presence of a primarily informational masker. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1861598]

PACS numbers: 43.66.Dc, 43.66.Lj, 43.66.Pn [AK]

Pages: 2169–2180

I. INTRODUCTION

The “cocktail party” environment (Cherry, 1953) is particularly difficult for many listeners with sensorineural hearing loss.¹ This environment contains multiple speech sources and often a listener faces the problem of understanding one talker while ignoring one or more others speaking at the same time. Cherry (1953) identified spatial separation between sound sources as one of several potential means by which listeners solve the cocktail party problem. A speech signal masked by speech is affected by at least two types of masking: *energetic* and *informational* (Freyman *et al.*, 1999, 2001; Brungart, 2001; Brungart *et al.*, 2001; Arbogast *et al.* 2002; see also earlier work by Carhart *et al.*, 1968). A noise, on the other hand, generally produces primarily energetic masking.

Energetic masking originates in the peripheral auditory system where the response of the auditory nerve to the masker does not change appreciably when the signal is added. Informational masking, on the other hand, originates at some point beyond the auditory periphery and occurs despite an adequate representation of the signal in the system. At the simplest level, informational masking is masking that cannot be accounted for by energetic masking, similar to what Carhart *et al.* (1968) referred to as “perceptual masking.” Informational masking is believed to occur due to listener uncertainty and/or to a high degree of similarity be-

tween the masker and the signal along one or more relevant stimulus dimensions (e.g., Watson, 1987; Neff and Green, 1987; Kidd *et al.*, 2002b; Durlach *et al.*, 2003b). There is currently considerable discussion about how best to define or compute informational masking and how to relate the results from different speech and nonspeech tasks (e.g., Durlach *et al.*, 2003a). For the purposes of the current study and for these stimuli and experimental task, informational masking is defined as masking beyond that which can be accounted for by energetic masking and is largely due, we believe, to similarity between the signal and masker, often causing them to be confused or attention to be directed to the wrong source.

In normal-hearing listeners the benefit of spatial separation of sources depends on the type of masking. Using non-speech stimuli, Kidd *et al.* (1998) found a significantly larger spatial release from an informational masker (5–40 dB) than from an energetic masker (less than 10 dB) for the identification of pure-tone frequency patterns. Freyman *et al.* (1999), using natural speech stimuli, found no spatial release from a noise masker when the precedence effect was used to create a *perceived* spatial separation between signal and masker. However, there was a 4- to 10-dB spatial release for a speech masker. They concluded that the release was from the informational portion of the masking produced by the speech masker and due to perceptual segregation of the two sources in space (Freyman *et al.*, 1999, 2001). However, natural speech signals and maskers played simultaneously have substantial spectral overlap and thus are likely to contain significant amounts of energetic masking. They also produce an amount of informational masking that is difficult to quantify or control. Therefore, Arbogast *et al.* (2002) used

^{a)}Portions of this research were presented at the Midwinter Meeting of the Association for Research in Otolaryngology, Daytona Beach, FL, January 2003. This research has been included as a portion of a doctoral dissertation by the first author.

^{b)}Electronic mail: tanya.arbogast@earthlink.net

processed speech stimuli in order to control, as best as possible, the ratio of informational to energetic masking. They measured the spatial release from masking for a 90° separation of signal and masker and found an 18-dB release from a primarily informational masker and a 7-dB release from a primarily energetic masker. The larger advantage for the informational masker was attributed to the perceptual effect of spatial separation, similar to the interpretation of Freyman *et al.* (1999, 2001). Hearing-impaired listeners generally obtain a spatial release from masking (noise or natural speech masking) that is smaller than normal-hearing listeners (e.g., Peissig and Kollmeier, 1997; Duquesnoy, 1983; Bronkhorst and Plomp, 1989; Hawley, 2000). However, it is not known if these listeners can use the perceptual effect of spatial separation to improve performance in the presence of a masker that is primarily informational.

Previous studies using pure-tone stimuli have found that hearing-impaired listeners can use segregation cues to obtain a release from informational masking, but the release is either smaller or requires a more obvious cue than for normal-hearing listeners [Grose and Hall (1996) for a melody recognition task; Kidd *et al.* (2002a) for detection of a spectrotemporally coherent tone in a spectrotemporally incoherent masker]. It is not known whether listeners with hearing loss can use spatial separation cues to overcome informational masking. Sound sources in realistic environments are often located in different places, affording potential cues for perceptually segregating the sounds they produce. Therefore, it is of considerable interest to determine how well listeners with hearing loss can use spatial cues to reduce informational masking.

The current study investigated the effect of informational versus energetic masking and normal versus impaired hearing on spatial release from masking. Listeners with normal hearing were included to provide age-matched controls. This study used the same stimuli and procedures as Arbogast *et al.* (2002). These stimuli and procedures have the advantage of retaining a high degree of speech intelligibility in a straightforward task that is perhaps less abstract than some nonspeech detection/discrimination tasks while providing better isolation of informational masking than is possible using natural speech. The speech stimuli in the current experiments were a modified version of cochlear implant simulation speech (Shannon *et al.*, 1995) processed in a way that greatly reduces spectral overlap between signal and masker, providing greater isolation of informational masking. In addition to the masker designed to produce primarily informational masking, two other maskers were included. One masker was a control for any potential energetic masking caused by the informational masker. The other was constructed to maximize energetic masking and minimize informational masking, and was included in order to compare the spatial release from masking obtained for primarily informational and primarily energetic maskers. Therefore, the effect of spatial separation was examined as a function of three types of maskers that generate different amounts of masking and which are assumed to originate primarily in either the peripheral (energetic) or the central (informational) portions of the auditory system. The current methods and procedures

TABLE I. Demographic and audiologic data for the hearing-impaired listeners. PTA is the pure-tone average of 500, 1000, and 2000 Hz, in dB HL. WRS (R/L)=word recognition score (right/left). HA=hearing aid.

	Age	Sex	Length of loss (years)	PTA	WRS (R/L)	Etiology	HA use
HI1	21	F	5	36	100/96	probably genetic	none
HI2	74	M	10	28	88/92	unknown/age/noise?	none
HI3	47	F	7	33	92/96	ototoxicity—vancomycin	bilateral, occasionally, 1½ years
HI4	68	F	5	33	92/88	unknown	none
HI5	54	F	10	39	100/96	probably genetic	bilateral, frequently, 3 years
HI6	19	F	19	40	80/92	congenital/pre-natal	bilateral, frequently, 1½ years
HI7	26	F	18	37	92/96	unknown/genetic?	none
HI8	37	F	37	43	96/92	genetic	bilateral, frequently, 32 years
HI9	67	F	60	48	72/80	unknown/genetic?	bilateral, frequently, 35 years
HI10	79	F	several years	48	80/88	unknown/age?	bilateral, occasional, 1 year

were nearly identical to Arbogast *et al.* (2002) and therefore will only be described briefly here.

II. METHODS

A. Listeners

Two groups of ten listeners were tested. Listeners received monetary compensation for their participation. The normal-hearing (NH) group had audiometric pure-tone thresholds of 20 dB HL or better in each ear for frequencies between 250 and 6000 Hz. In order to ensure a similar age range in each listener group, each NH listener was roughly matched in age (within a decade) to a HI listener. NH listeners ranged in age from 19 to 77 years. Hearing-impaired listeners ranged in age from 19 to 79 years. The hearing-impaired (HI) group was composed of listeners with bilateral, symmetrical sensorineural hearing loss of probable cochlear origin. Symmetrical loss was defined as pure-tone audiometric air-conduction thresholds within ± 10 dB in each ear at most audiometric frequencies between 250 and 6000 Hz. Table I details demographic and audiologic information about the HI listeners. The duration and etiology of hearing loss, as well as hearing aid use, was obtained from a listener questionnaire. Listeners were ordered from lowest to highest threshold in quiet for the speech stimuli used in this study.

The reported duration of hearing loss ranged from 3 to 60 years. The degree of loss ranged from mild to moderate and the configuration ranged from flat to gradually sloping. The mean pure-tone average (of 500, 1000 and 2000 Hz) was

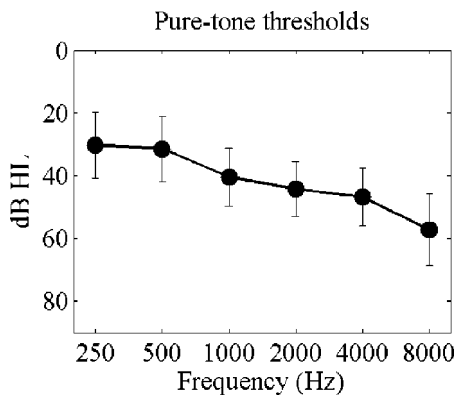


FIG. 1. Mean audiometric thresholds in dB HL for the hearing-impaired listeners. Error bars are ± 1 standard deviation.

37 dB HL. Word recognition scores ranged from 72% to 100% correct but were within $\pm 12\%$ between the ears of any individual listener, using standard, monosyllabic materials. Six of the ten HI listeners wore hearing aids (all bilaterally; testing was unaided). Figure 1 plots mean audiometric thresholds (obtained using standard clinical procedures) in dB HL for the HI group.

B. Stimuli

The stimuli were the same as those used by Arbogast *et al.* (2002) and are described in detail in that article. A brief description is given here. The four male talkers of the coordinate response measure (CRM) corpus of sentences (Bolia *et al.*, 2000) were used. Each sentence had the structure: “Ready [callsign] go to [color] [number] now.” Signal sentences had the callsign “Baron” and each was paired with a randomly chosen masker sentence with a different callsign, color, number, and talker. Each sentence was digitally pre-processed to produce a set of 15 pure tones modulated by the speech envelope in the $\frac{1}{3}$ -octave band centered at each frequency. Center frequencies were evenly spaced on a logarithmic scale from 215 to 4891 Hz. For details of the preprocessing, see Arbogast *et al.* (2002). Signal sentences were generated by combining eight (out of 15) randomly chosen bands. Each masker sentence was used to create three different types of maskers (described below). Signal and masker sentence onset was simultaneous.

1. Different-band sentence (DBS) masker

The different-band sentence (DBS) masker was generated by combining six randomly chosen bands excluding the eight bands used for the signal. Therefore, each signal/DBS masker pair contained mutually exclusive bands. The top panel of Fig. 2 shows the magnitude spectra of a signal/DBS masker pair. The signal is gray and the masker is black. The DBS masker was intended to produce primarily informational masking because it did not overlap the signal in frequency and was similar to the signal sentence in terms of intelligibility (Brungart *et al.*, 2004), overall sound quality due to the pure-tone carriers, and use of the same speech corpus for signal and masker.

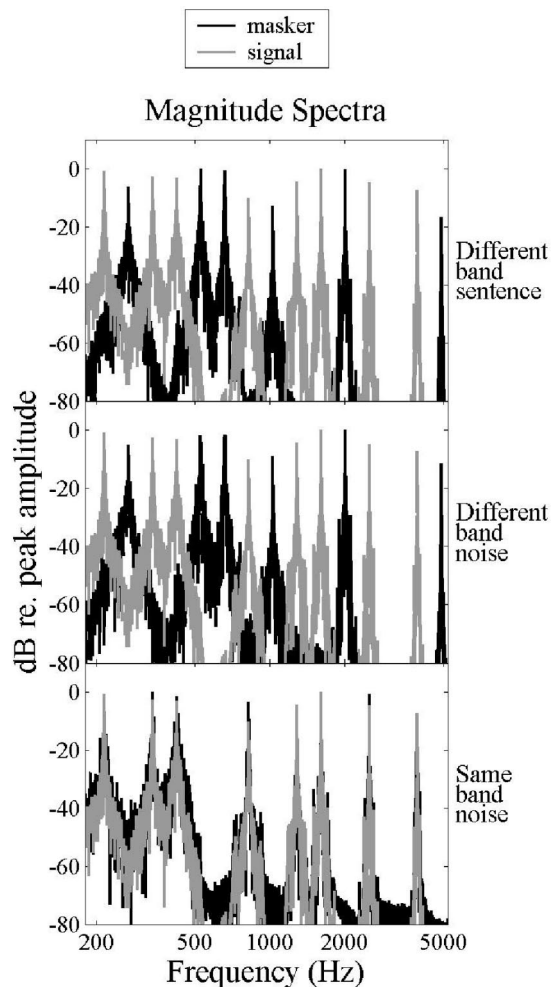


FIG. 2. Magnitude spectra for a sample signal sentence and the three types of processing for the masker sentence. The same signal sentence is plotted in gray in each panel. The black indicates the masker sentence for the different-band sentence masker (top panel), the different-band noise masker (middle panel), and the same-band noise masker (bottom panel).

2. Different-band noise (DBN) masker

The long-term complex spectrum of each DBS masker, described above, was used to shape the spectrum of a random, broadband Gaussian noise and the result was inverse fast Fourier transformed. This masker was intended to produce the same minimal amount of energetic masking as the DBS masker, but also to produce minimal informational masking because it was unintelligible and qualitatively dissimilar from the signal sentence.² This masker was included as a control for energetic masking in the DBS masker. The magnitude spectra of a DBN/signal sentence pair are shown in the middle panel of Fig. 2. The envelope of the DBN masker contains amplitude modulation because the stimulus is essentially the sum of multiple narrow-band noises.³

3. Same-band noise (SBN) masker

This masker was comprised of eight bands of noise that exactly overlapped the eight bands of the signal with which it was paired. The long-term complex spectrum of the eight-band masker sentence was used to shape the spectrum of a broadband Gaussian noise, and the result was inverse fast

Fourier transformed. The magnitude spectra of a SBN/signal sentence pair are shown in the lower panel of Fig. 2. Energetic masking was maximized because the magnitude spectra of signal and masker overlapped. Informational masking was minimized because the masker had no intelligibility and, as with the DBN masker, was qualitatively dissimilar to the signal sentence. The envelope of the SBN masker is similar to that of the DBN masker.

C. Procedures

The experiment took place in a soundfield, located within a single-walled, 12×13 ft², IAC sound booth. When two stimuli were present, they were played simultaneously via two Pioneer S-DF3-K speakers located 5 ft from the listener, one at 0° azimuth and the other at 90° azimuth to the right in the horizontal plane approximately level with the ears. The signal was always played from the speaker located at 0° azimuth. Stimuli were played through Tucker-Davis Technologies (TDT) hardware. The signal and the masker sentences were played through separate channels. Stimuli were converted at a rate of 50 kHz via a 16-bit, 8-channel D/A converter (DA8), low-pass filtered at 20 kHz (FT-6), and attenuated (PA-4). The signal was sent through a programmable switch (SS-1) to route it to the desired signal speaker. The signal and masker were sent to mixers where they were summed if presented to the same speaker. The stimuli were passed through power amplifiers (Tascam) and then to one or two speakers in the soundfield, depending on the experimental condition.

The task was one-interval forced-choice with two responses required on each trial. The first response (color) had four alternatives (blue, red, green, white), while the second response (number) had eight alternatives (1–8). Signal-alone thresholds were obtained using 30-trial blocks of a one-up, one-down adaptive procedure that estimates the 50% correct point on the psychometric function (Levitt, 1971). Masked data were collected with the method of constant stimuli. Four signal levels were chosen for each listener and condition individually in order to estimate the psychometric function. The aim was to produce performance ranging from just above chance to nearly 100% correct. The exact signal levels and level intervals varied depending on the listener, the condition, and the early psychometric function results. Signal levels were mixed randomly within every 50-trial block. On each trial, the listener was required to identify the color and number from the sentence with the call sign “Baron” by entering values on a computer keyboard. Response feedback was provided after each trial.

Each of the three types of maskers was tested in two spatial separation conditions, 0° and 90°, for a total of six masked conditions. In addition, one signal-only condition was tested. The signal sentence was always played at the 0° location and the masker, when present, was played at either the 0° or the 90° location. Each block of trials contained only one spatial separation condition and one masker type. The masker level was set for each listener to approximately 40 dB SL (above signal-alone threshold). However, if comfort or equipment limits were lower, then the highest possible masker level was used. The equipment limit was 88 dB SPL,

but in order to allow for a +10 dB S/M (necessary in most cases to define the upper portion of the psychometric function), the highest masker level permitted was 78 dB SPL. Two blocks of signal-alone identification thresholds were collected first. Masked data were collected in sets of two blocks, each of the same masker/spatial separation condition. Sets of blocks containing each masker and spatial separation condition were alternated to balance any learning effects equally across all conditions. A minimum of 100 trials per signal level was obtained for each spatial separation and masker type. Data were collected within four sessions of 2 h each. Listeners were instructed to keep the head straight, facing the 0° location throughout stimulus presentation.

Training included a 50-trial block of practice of the signal-alone condition. The signal sentences were presented at a comfortable level from the 0° speaker. After achieving a minimum of 90% correct they completed four 50-trial practice blocks of the masked condition with the DBS masker only. This practice included two blocks each for the two spatial separation conditions. The S/M was fixed at +15 dB for the first block and +10 dB for the second block.

Each response was considered correct only if both the color and number of the signal sentence were identified accurately. The psychometric functions were fit by the Nelder-Mead simplex method (FMINS function in MATLAB) with a logistic of the form

$$p(c) = \alpha + ((1.0 - \alpha) / (1.0 + e^{-k(x-m)})), \quad (1)$$

where $\alpha=0.03125$ (chance performance for 1 out of 32), k is the slope of the function, m is the level of the midpoint (about 51.5% correct), and x is the signal sentence level.

III. RESULTS

The amount of spatial release (SR) from masking was calculated as the S/M required for 51% correct performance for the 0° condition minus that for the 90° condition. The amount of informational masking created by the DBS masker was estimated by taking the difference between the midpoints of the psychometric functions for the DBS and DBN maskers presented at 0°. This estimate was based on the type of masking each of the two maskers was intended to produce as discussed above. The purpose of the estimate of informational masking in the current study was to provide a means of comparing performance between listeners and groups.

All listeners were able to perform the task with greater than 90% accuracy in quiet within the first one to two practice blocks. Practice performance for the signal presented with the informational masker was generally good, ranging from 76% to 100% correct, depending on spatial separation and S/M. Masker SLs for the NH group ranged from 33.5 to 43.9 dB with a mean of 38.4 dB. Masker SLs for the HI group ranged from 21.5 to 38.5 dB with a mean of 32.5 dB.

All psychometric functions were orderly and monotonic. Overall, the results were consistent with those from the earlier study (Arbogast *et al.*, 2002), which did not exhibit performance plateaus or dips at S/Ms just below 0 dB. The logistic fits to both the NH and HI data were very good. The

TABLE II. Signal-to-masker ratio at m for the normal-hearing listeners for the two spatial separation conditions. The results for each masker are given in separate subtables. Quiet threshold (Quiet) in dB SPL and masker sensation level (SL) in dB, as well as variance accounted for by the logistic fit (var), are also included. The final column gives the difference between spatial conditions for m .

Listener	Quiet	SL	0° separation		90° separation		0°–90° m
			S/M	var	S/M	var	
A. Different-band sentence masker							
1	16.5	33.5	-17.4	>0.99	-23.7	>0.99	6.3
2	21.9	38.1	4.3	0.98	-14.8	0.98	19.1
3	18.8	36.2	-0.1	>0.99	-10.0	0.96	10.0
4	16.7	38.3	1.1	>0.99	-14.9	0.97	16.0
5	18.0	42.0	-4.0	0.96	-19.7	>0.99	15.7
6	10.0	40.0	-0.9	0.98	-17.9	>0.99	17.0
7	6.1	43.9	-7.8	0.98	-22.8	0.99	15.0
8	11.4	38.6	-4.6	0.99	-26.0	0.95	21.4
9	13.1	36.9	-1.0	0.98	-14.7	>0.99	13.6
10	18.4	36.6	-1.3	0.99	-20.1	0.98	18.7
Average	15.1	38.4	-3.2	...	-18.5	...	15.3
s.d.	4.8	3.0	6.0	...	4.9	...	4.5
B. Different-band noise masker							
1	16.5	33.5	-27.4	0.99	-28.9	0.99	1.5
2	21.9	38.1	-22.3	0.98	-28.4	>0.99	6.1
3	18.8	36.2	-22.8	>0.99	-26.2	>0.99	3.4
4	16.7	38.3	-23.7	>0.99	-25.2	0.99	1.6
5	18.0	42.0	-23.5	>0.99	-25.7	>0.99	2.2
6	10.0	40.0	-25.6	>0.99	-30.0	>0.99	4.4
7	6.1	43.9	-28.8	>0.99	-32.4	0.99	3.5
8	11.4	38.6	-29.4	0.99	-33.6	0.92	4.1
9	13.1	36.9	-25.5	>0.99	-27.8	>0.99	2.3
10	18.4	36.6	-23.1	>0.99	-26.8	>0.99	3.7
Average	15.1	38.4	-25.2	...	-28.5	...	3.3
s.d.	4.8	3.0	2.6	...	2.8	...	1.4
C. Same-band noise masker							
1	16.5	33.5	-1.7	0.99	-7.3	0.99	5.6
2	21.9	38.1	1.9	0.99	-5.9	>0.99	7.8
3	18.8	36.2	-0.9	0.99	-5.0	0.99	4.1
4	16.7	38.3	2.4	0.99	-2.6	0.99	5.1
5	18.0	42.0	-0.2	>0.99	-4.1	>0.99	3.9
6	10.0	40.0	0.0	>0.99	-6.1	>0.99	6.2
7	6.1	43.9	-0.2	0.99	-8.9	0.99	8.7
8	11.4	38.6	-0.7	>0.99	-8.5	0.99	7.8
9	13.1	36.9	-0.6	>0.99	-4.9	0.98	4.3
10	18.4	36.6	-0.1	0.99	-5.8	0.99	5.7
Average	15.1	38.4	0.0	...	-5.9	...	5.9
s.d.	4.8	3.0	1.3	...	1.9	...	1.7

variance accounted for by the fit in each instance was greater than 92%, and in most cases it was 98% or better. Tables II and III list the S/M at the midpoint (m) from the fitted psychometric functions for NH and HI listeners, respectively. HI listeners are numbered in the same manner as in Table I (ordered by quiet threshold). Each NH listener was matched with a HI listener by age, and then given the same number as their HI counterpart.

A. Signal-to-masker ratios

Figure 3 plots the mean S/M at the midpoint of the psychometric function for the two listener groups and spatial separation conditions. Each panel is for a different masker.

The figure reveals that S/M varied with masker type, listener group, and spatial separation condition. The signifi-

cance of these variations is discussed in more detail below in the context of a statistical analysis of the data.

1. Factors influencing S/M

The S/M data were submitted to an ANOVA with one between-subjects factor (group—NH/HI) and two within-subjects factors (masker and spatial separation). All three main effects were significant: masker [$F(2,36)=330.6$, $p<0.001$], spatial separation [$F(1,18)=237.6$, $p<0.001$], and group [$F(1,18)=20.6$, $p<0.001$]. *Posthoc* simple contrasts revealed that the average S/M ratio for each masker was significantly different from the others. The interaction of masker and spatial separation was significant [$F(2,36)=76.1$, $p<0.001$] as well as the interaction of masker and

TABLE III. Same as Table II, but for the hearing-impaired listeners.

Listener	Quiet	SL	0° separation		90° separation		0°-90° <i>m</i>
			S/M	var	S/M	var	
A. Different-band sentence masker							
1	36.7	38.3	-3.3	>0.99	-16.2	0.98	12.9
2	37.0	31.0	1.9	0.98	-7.4	0.98	9.2
3	39.5	38.5	1.2	0.99	-10.9	0.99	12.0
4	40.4	37.6	2.4	0.95	-7.3	0.99	9.7
5	43.3	34.7	2.1	0.99	-6.3	0.99	8.4
6	43.5	34.5	-1.9	0.98	-15.2	0.99	13.3
7	44.0	34.0	-2.2	0.99	-14.8	>0.99	12.6
8	45.7	32.3	-2.0	0.99	-12.5	0.96	10.5
9	55.2	22.8	3.6	0.99	-0.4	>0.99	4.0
10	56.5	21.5	3.7	0.98	1.6	0.99	2.0
Average	44.2	32.5	0.5	...	-8.9	...	9.5
s.d.	6.8	6.0	2.6	...	6.1	...	3.8
B. Different-band noise masker							
1	36.7	38.3	-18.7	0.98	-22.4	0.99	3.7
2	37.0	31.0	-15.4	>0.99	-22.5	0.99	7.0
3	39.5	38.5	-12.0	>0.99	-15.6	>0.99	3.5
4	40.4	37.6	-9.0	>0.99	-17.2	0.99	8.2
5	43.3	34.7	-7.8	>0.99	-12.3	0.99	4.5
6	43.5	34.5	-16.7	0.99	-18.3	>0.99	1.6
7	44.0	34.0	-15.7	0.99	-18.6	>0.99	2.9
8	45.7	32.3	-13.4	>0.99	-17.7	>0.99	4.3
9	55.2	22.8	-5.8	0.99	-7.2	>0.99	1.4
10	56.5	21.5	-3.5	0.98	-6.8	0.99	3.3
Average	44.2	32.5	-11.8	...	-15.9	...	4.0
s.d.	6.8	6.0	5.1	...	5.5	...	2.1
C. Same-band noise masker							
1	36.7	38.3	-0.8	>0.99	-9.8	>0.99	9.0
2	37.0	31.0	1.3	>0.99	-7.7	0.99	9.0
3	39.5	38.5	0.6	0.99	-5.4	0.99	6.0
4	40.4	37.6	2.3	0.97	-3.4	0.99	5.7
5	43.3	34.7	1.6	>0.99	-4.7	>0.99	6.2
6	43.5	34.5	-1.1	>0.99	-6.9	0.99	5.8
7	44.0	34.0	-0.6	>0.99	-9.0	>0.99	8.4
8	45.7	32.3	-1.4	>0.99	-8.8	>0.99	7.4
9	55.2	22.8	1.5	0.99	-3.0	0.99	4.4
10	56.5	21.5	1.4	0.99	-0.5	0.96	1.9
Average	44.2	32.5	0.5	...	-5.9	...	6.4
s.d.	6.8	6.0	1.3	...	3.0	...	2.2

group [$F(2,36)=41.1, p<0.001$]. The three-way interaction between masker, group, and spatial separation was also significant [$F(2,36)=13.1, p<0.001$].

Three important points regarding the S/M data were confirmed by the ANOVA results. First, the masker by group interaction revealed that S/M was greater for the HI group than the NH group for the DBS and DBN maskers, but not for the SBN masker. Perhaps most importantly, for the DBN masker the HI group S/M was 13 dB greater than for the NH group. Recalling that the DBN masker estimates energetic masking in the DBS masker, this indicates that the HI group experienced significantly more energetic masking than the NH group for both the DBN and DBS maskers. Second, the main effect of masker and posthoc tests revealed that the DBS masker resulted in a significantly greater S/M than the DBN masker, suggesting that a substantial amount of informational masking was produced by the DBS masker. At 0° the DBS masker caused an additional 22.0 dB of masking in the NH group and an additional 12.3 dB of masking in the HI

group. Third, SR depends on both the type of masker and the listener group.

This three-way interaction is illustrated in Fig. 4 which plots the SR obtained by each group for each of the three maskers. To investigate this interaction further, three one-way ANOVAs were performed on the SR data, one for each masker type, with group as the single between-subjects factor. The results revealed that the SR for the DBS masker was significantly greater for the NH group (15.3 dB) than for the HI group (9.5 dB) [$F(1,18)=9.9, p<0.01$]. However, the SR for the other two maskers was similar for the two groups (3.3 and 4.0 for the DBN [$F(1,18)=0.9, p=0.37$]; 5.9 and 6.4 for the SBN [$F(1,18)=0.3, p=0.59$], for NH and HI, respectively).

Although the HI group obtained a smaller SR than the NH group for the DBS masker, the HI group's advantage was still larger for this masker than it was for the SBN masker (by about 3 dB on average). However, this 3-dB mean difference in SR between the DBS and SBN maskers for the HI

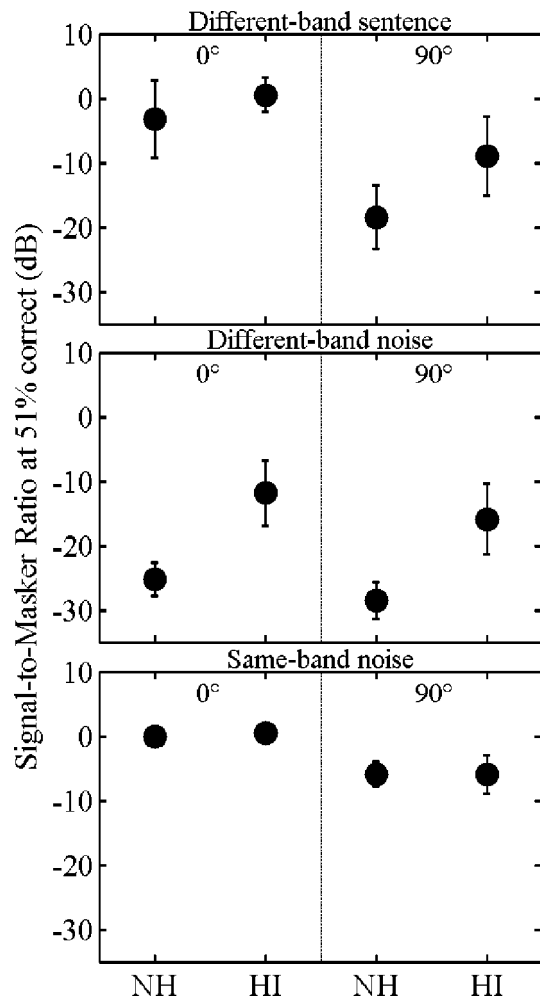


FIG. 3. Group mean signal-to-masker ratios at the 51% correct point on the psychometric function plotted for each listener group (NH/HI). The top panel is for the DBS masker, the middle is for the DBN masker, and the bottom panel is for the SBN masker. Data for the 0° separation condition are plotted in the left half of each panel, and the 90° separation data are plotted in the right half of each panel. The error bars are ± 1 standard deviation.

group is small relative to the nearly 10-dB difference found for the NH group.

2. Factors related to spatial release from informational masking

The NH and HI groups differed in SR for the DBS masker only. Therefore, a correlational analysis was performed which focused on the SR from this masker. One of the factors investigated was an estimate of the amount of informational masking produced by the DBS masker. The DBS masker produced a mean of 22.0 dB of informational masking in the NH group, but a mean of only 12.3 dB of informational masking in the HI group. The majority of the difference between groups can be accounted for by the difference in S/M for the DBN masker (13 dB greater for the HI group than the NH group).

For the NH group, Pearson correlation coefficients were not significant ($p > 0.05$) between SR for the DBS masker and the factors of age ($r = 0.38$), quiet threshold ($r = -0.09$), or masker SL ($r = 0.45$). SR in the NH group was significantly correlated with the estimated amount of infor-

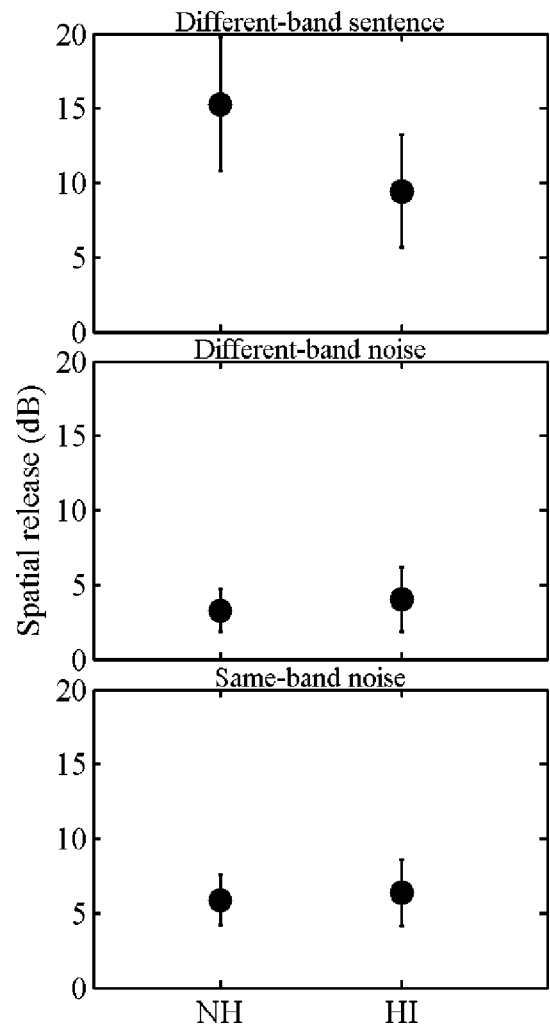


FIG. 4. Spatial release from masking for the two listener groups. The top panel is for the DBS masker, the middle panel is for the DBN masker, and the lower panel is for the SBN masker. The data are group means and the error bars are ± 1 standard deviation.

mational masking ($r = 0.74$, $p < 0.02$). However, a single outlier (NH1) appeared to be responsible for this significant relation. When the outlier was excluded, the correlation was no longer significant ($r = 0.3$, $p = 0.39$).

For the HI group, SR was significantly correlated with the amount of informational masking ($r = 0.76$, $p < 0.01$). In addition, SR was correlated with age ($r = -0.81$, $p < 0.01$), quiet threshold ($r = -0.81$, $p < 0.01$), masker SL ($r = 0.88$, $p < 0.01$), and the mean word recognition score for the two ears ($r = 0.65$, $p < 0.05$). SR was not significantly related to hearing aid use ($r = -0.4$, $p = 0.29$), length of hearing aid use ($r = -0.2$, $p = 0.69$), or length of hearing loss ($r = -0.4$, $p = 0.32$). Age and quiet threshold were not significantly correlated with each other ($r = 0.39$, $p = 0.27$) and, therefore, these two factors may have influenced SR independently. However, quiet threshold and masker SL were significantly correlated with each other ($r = -0.88$, $p < 0.01$), as expected because the masker level was mainly determined by quiet threshold. However, the existence of this correlation means that the relation between SR and quiet threshold may have been a by-product of masker SL. In order to tease these variables apart, a stepwise regression

analysis was performed with SR from the DBS masker as the dependent variable and amount of informational masking, age, quiet threshold, masker SL, and mean word recognition score as the independent variables. The analysis revealed that the combination of amount of informational masking, masker SL, and age explained 97% of the variance in SR from the DBS masker ($p < 0.05$ for all three variables). Quiet threshold and mean word recognition score did not significantly increase the variance accounted for ($p > 0.10$). Type II partial correlations (variable added last) revealed that the amount of informational masking accounts for 6%, masker SL accounts for 12%, and age accounts for 10% of the variance in SR from the DBS masker.

3. The influence of masker sensation level on spatial release from informational masking

Overall, the HI group was tested at a lower masker SL (32.5 dB) than the NH group (38.4 dB). While the mean difference in sensation level was not large, individual sensation levels ranged from 21.5 dB up to 43.9 dB. Therefore, the analysis above was repeated after excluding two from each group. The masker SL for each of these four listeners was more than 1 standard deviation from the mean masker SL of all 20 listeners. NH5, NH7, HI9, and HI10 were excluded. After excluding these listeners, group mean masker sensation levels were 37.3 dB for the NH group and 35.1 dB for the HI group.

SR for the DBS masker remained larger for the revised NH group (15.3 dB) than for the revised HI group (11.1 dB). However, SR for the revised HI group was larger than for the original HI group (9.5 dB). The small increase in SR was mainly due to better performance in the 90° condition. An independent-sample *t*-test revealed that the SR for the DBS masker was not significantly different between the revised NH and HI listener groups [$t(8.9) = 2.2, p = 0.06$].

Four of the NH listeners (NH2, NH5, NH8, and NH10) were retested using a lower masker sensation level. The mean masker SL was 38.8 dB for their first run through the experiment. The second time through the experiment, the mean masker SL was 25.8 dB. SR averaged 18.7 dB for the higher masker SL and 12.3 dB for the lower masker SL. This difference just missed reaching significance on a paired-sample *t*-test [$t(3) = 2.7, p = 0.07$], perhaps due to the small number of listeners. However, in all four cases, SR from the DBS masker was smaller for the lower masker SL than for the higher masker SL.

B. Errors

The proportion of incorrect responses matching the masker sentence was calculated excluding data at the lowest and highest signal levels tested for each listener. This left two signal levels in the middle of the psychometric function from which to analyze the error patterns. Very low S/Ms were excluded because listeners may have purposely responded with some other color/number combination than that heard clearly from the masker sentence in order to increase their chance of guessing correctly. Very high S/Ms were excluded because few errors occurred.

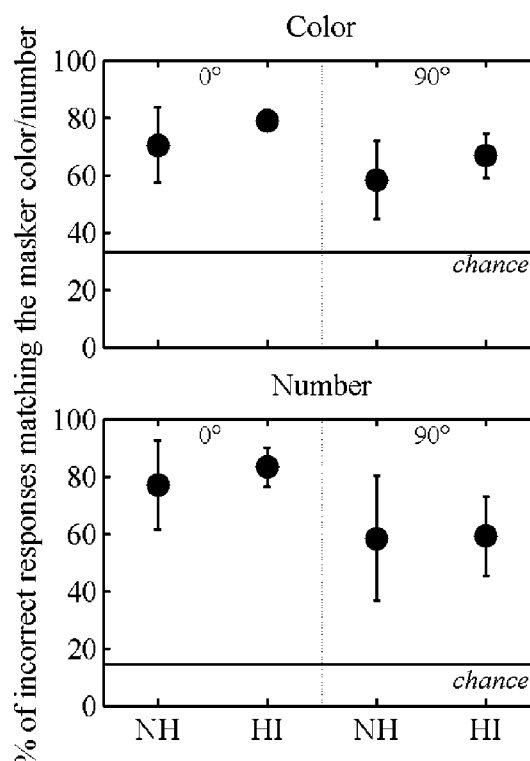


FIG. 5. Percent of incorrect responses that matched the DBS masker color (top panel) or number (lower panel) for the two listener groups. The 0° data are in the left half of each panel and the 90° data are in the right half. The data are group means and the error bars are ± 1 standard deviation. The percent of errors expected by chance is given by the solid line in each panel (33.3% for color and 14.3% for number).

Figure 5 shows the mean percent of errors that matched the DBS masker color and number for the two spatial separation conditions and listener groups. The top panel is for the color response and the lower panel is for the number response. The percent of errors expected for a random distribution of incorrect responses is indicated by the solid line in each panel (33.3% for color; 14.3% for number).

The mean percents of errors from the masker for the NH group in the 0° condition were 70% and 77% for color and number, respectively; for the HI group they were 79% and 83%, respectively. This indicates that listeners in both groups often confused the signal and masker sentences when they were played from the same location. In the 90° condition, the percents of errors from the masker for the NH group were 58% each for color and number; and for the HI group they were 67% and 59% for color and number, respectively. Therefore, both groups of listeners confused the signal and masker sentences less often in the 90° condition than in the 0° condition. A repeated-measures ANOVA with group as the between-subjects factor and spatial separation as the within-subjects factor confirmed a significant main effect of spatial separation for both color [$F(1,18) = 32.37, p < 0.001$] and number [$F(1,18) = 28.86, p < 0.001$]. However, the main effect of group just missed significance for the color response [$F(1,18) = 4.3, p = 0.053$], and was not significant for the number response [$F(1,18) = 0.39, p = 0.54$]. The interaction of spatial separation and group was also not significant for color [$F(1,18) \sim 0, p = 0.99$] or number [$F(1,18) = 0.50, p = 0.49$].

IV. DISCUSSION

A. Informational and energetic masking

Consistent with Arbogast *et al.* (2002), the DBS masker appears to have produced the intended type of masking in both normal-hearing and hearing-impaired listeners. The DBS masker caused 22 dB of informational masking for the normal-hearing group and 12 dB of informational masking for the hearing-impaired group. In a different group of normal-hearing listeners Arbogast *et al.* (2002) also found 22 dB of informational masking. The large proportion of confusions, in both listener groups, between the signal and DBS masker also supports the interpretation that a significant amount of informational masking was present (Brungart, 2001; Brungart *et al.*, 2001; Arbogast *et al.*, 2002). For the 0° condition, both listener groups responded with the masker color and number an average of 75% and 80% of the time, respectively. In the previous group of normal-hearing listeners, Arbogast *et al.* (2002) found the same.

However, the DBS masker produced less informational masking in the hearing-impaired group than in the normal-hearing group. This may be due to greater energetic masking in hearing-impaired than normal-hearing listeners as measured by the DBN masker,⁴ probably as a result of wider auditory filters (e.g., Glasberg and Moore, 1986; Leek and Summers, 1993). The greater amount of energetic masking for the DBN masker for the hearing-impaired group suggests that they will experience a greater amount of energetic masking for the DBS masker also. There are at least two reasons why more energetic masking in the DBS masker might lead to less informational masking in that same masker. One reason is a ceiling effect. Given the baseline S/M due to energetic masking (DBN masker) for the hearing-impaired group (−12 dB), if the DBS masker produced the same amount of informational masking as in normal-hearing listeners (22 dB), one would expect the DBS masker to result in a mean midpoint S/M of about +10 dB for the hearing-impaired group (−12 dB +22 dB of informational masking = +10 dB). However, the actual mean for this group was close to 0 dB S/M. The hearing-impaired listeners may have been able to segregate the signal and masker sentences when the signal sentence was louder than the masker sentence. Therefore, there may have been a critical point just beyond 0 dB S/M for which the available/effective cues changed and the task became relatively easy. Most of the normal-hearing listeners did not reach that ceiling because of the lower S/M for the DBN masker. Brungart (2001), using the (unprocessed) CRM materials and test procedures, reported that performance in speech on speech masking for two different same-sex talkers (one target, one masker) was at about 60% correct at 0 dB S/M. Increasing the level of the target by only 6 dB increased performance to nearly 90% correct, consistent with the proposition that a positive S/M drives performance rapidly toward asymptote.

It is also possible that within a single masker, informational masking decreases when energetic masking increases. Kidd *et al.* (2003) measured harmonicity discrimination of a multi-tone complex in the presence of multi-tone informational maskers. The binaural advantage decreased as the

number of components in the informational masker increased. The interpretation was that increasing the number of components resulted in an increase in energetic masking because masker components were more likely to fall close in frequency to signal components. The increase in energetic masking caused a decrease in informational masking, resulting in a reduced binaural advantage. Similarly, other studies have reported a decrease in informational masking when energetic masking was increased by increasing the number of tones in a multi-tone masker for the task of detecting a pure-tone signal (e.g., Neff and Green, 1987; Oh and Lutfi, 1998).

B. Spatial release from masking

1. DBN masker

The spatial release from the DBN masker averaged 3.7 dB for all 20 listeners, and was not significantly different for the two listener groups. This is very similar to the 3.6-dB release from this same masker found by Arbogast *et al.* (2002) in normal-hearing listeners. It is difficult to compare this result to that of previous research because the masking was off-frequency and mainly energetic. However, one might expect a smaller release from this masker than from an on-frequency energetic masker based on the results of Zwicker and Henning (1984). They found that the masking level difference decreased significantly when the signal and masker did not overlap in frequency than when they were spectrally matched.

2. SBN masker

The mean spatial release from the SBN masker was 6.1 dB for all listeners, and was also not significantly different for the two listener groups. This value is consistent with that found by Arbogast *et al.* (2002) for the same masker (6.9 dB). Similarly, previous research has found spatial release from energetic masking in the range of 5 to 10 dB (Hawley, 2000; Duquesnoy, 1983; Peissig and Kollmeier, 1997; Bronkhorst and Plomp, 1988, 1989, 1990, 1992; Freyman *et al.*, 1999; Zurek, 1993) for a 90° separation between signal and masker. Other studies of hearing-impaired listeners have found mean spatial releases from energetic masking for suprathreshold speech recognition that are somewhat smaller than that found for normal-hearing listeners in the same study (Peissig and Kollmeier, 1997; Duquesnoy, 1983; Bronkhorst and Plomp, 1989). However, similar to the current research, two studies have found that the hearing-impaired group obtained nearly the same release (within 1.5–2 dB) as the normal-hearing group (Bronkhorst and Plomp, 1990; 1992).

3. DBS masker

The mean spatial release from the DBS masker was significantly larger for the normal-hearing group (15.3 dB) than for the hearing-impaired group (9.5 dB). However, for both groups, the spatial release from the DBS masker was significantly larger than the spatial release from either of the other two maskers. The size of the effect for the normal-hearing listeners was smaller than the 18.4-dB release found by Arbogast *et al.* (2002) for the same masker. This difference

may be attributed to the number and mix of listeners tested. The previous study tested four listeners and the range of spatial release produced was only 4 dB. All four listeners were fairly similar to each other for this task. In the current study, the range of spatial release was 15 dB. There were two listeners in particular (NH1 and NH3) with fairly small spatial releases. NH1 performed extremely well for the 0° condition and therefore could not improve by a large amount for the 90° condition. NH3, for some unknown reason, did not perform as well as the other normal-hearing listeners for the 90° condition.

Other studies have also found significant differences in spatial release between informational and energetic maskers. Kidd *et al.* (1998), using pure-tone stimuli, and Freyman *et al.* (1999), using natural speech stimuli, found that the spatial release from a masker producing a high proportion of informational masking was significantly larger than from a masker producing a high proportion of energetic masking. In Kidd *et al.* (1998), the release was 0 to 20 dB larger than the corresponding release from an energetic masker, depending on the listener and the signal frequency. In Freyman *et al.* (1999), the release was about 6 dB larger than from an energetic masker. Therefore, the larger spatial release found in the current study is consistent with the presence of a significant amount of informational masking. The expected spatial release due to binaural interaction and headshadow effects for the DBS masker can be estimated from the DBN masker, which had the same long-term magnitude spectrum as the DBS masker, but minimal informational masking. The release for this control masker was only 3.7 dB. Therefore 5 dB (for the hearing-impaired listeners) to 11 dB (for the normal-hearing listeners) of the spatial effect was beyond that expected for a primarily energetic masker with the same long-term magnitude spectrum as the DBS masker, and therefore most likely attributable to release from informational masking.

It is difficult to compare the results for the DBS masker to other speech research because, in addition to using natural speech stimuli (for which the ratio of informational to energetic masking likely is smaller), stimuli were configured in a way that inadvertently reduced informational masking (reduced stimulus uncertainty by holding the masker sentence constant and/or used a strong segregation cue such as male versus female) in order for the listener to be able to label the signal sentence. Therefore, energetic masking probably played the dominant role in much of this work and the results are more comparable to those found for the SBN masker of the current study. For normal-hearing listeners, Hawley *et al.* (2004) did find a larger spatial release (10–12 dB) from two- and three-talker maskers (where all talkers were the same as the signal talker) than from a two- or three-talker speech-shaped noise or speech-envelope-modulated speech-shaped noise (6–7 dB). However, these conditions were not tested in hearing-impaired listeners.

The results of previous research using pure-tone stimuli have suggested that hearing-impaired listeners do not benefit as much as normal-hearing listeners from segregation of a signal from an informational masker using cues other than spatial separation (cf. Kidd *et al.*, 2002a; Grose and Hall,

1996). Although the results of the current study appear to agree, two factors complicate this interpretation. One factor is that not all listeners in this study were tested at approximately the same masker sensation level and this appears to have at least partially influenced the results, as illustrated by the results of the reanalysis taking sensation level into account. The second factor, which may be related, is that the DBS masker did not produce as much informational masking in the hearing-impaired listeners as in the normal-hearing listeners. If both of these factors were carefully controlled, the difference between groups might very well diminish. Controlling for sensation level is entirely feasible, but controlling for informational masking may be more of a challenge.

Sensation level may have an impact on spatial release because both may be related to the amount of informational masking produced. Alexander and Lutfi (2003) found that informational masking decreased with decreasing masker sensation level in pure-tone stimuli. In addition, it is logical that as informational masking decreases, spatial release from informational masking will also decrease because there is less masking from which to be released. However, the reason for the relation between sensation level and amount of informational masking is not clear. The current research also found other factors relating to spatial release from the DBS masker in the hearing-impaired listeners, including (but not limited to) quiet threshold, age, and word recognition score.

Spatial release from informational masking was also reflected in the errors made by both groups of listeners. When the DBS masker was moved to the 90° position, the percents of incorrect responses matching the masker sentence dropped significantly by 12 and 20 percentage points for color and number responses, respectively. The decreased confusion between signal and masker implies a release from informational masking. This is similar to the finding of Brungart (2001) that when natural signal and masker sentences were segregated by voice fundamental frequency there was a decrease in the number of signal-masker confusions relative to when the two sentences were from the same talker or same-sex talkers.

The larger spatial advantage for the DBS masker than for the SBN masker implies that the perceptual effect of spatial separation between signal and masker (Freyman *et al.*, 1999, 2001; Arbogast *et al.*, 2002) can be used by both normal-hearing and hearing-impaired listeners to decrease informational masking. The ability to perceive that the signal and masker originate from separate locations in space allows the listener to label each sentence appropriately as “signal” or “masker,” effectively reducing confusion between the two sentences, thus reducing informational masking. This implies that the higher-level processes that allow for this perceptual effect are functioning effectively in listeners with peripheral hearing loss. This is consistent with the results of Noble *et al.* (1997), showing that both normal-hearing and hearing-impaired listeners can correctly indicate when two simultaneously presented, equal-level, suprathreshold sources (speech and noise) are spatially separated by 54° more than 85% of the time. Larger azimuthal separations were not tested, but one could speculate that performance would be at

least equal to, if not better than, 85% for a 90° separation. The current research, however, cannot determine if these higher-level processes in hearing-impaired listeners are functioning as well as they do in normal-hearing listeners. While the statistical results of the equal sensation level comparison suggest that they are, the hearing-impaired listeners still did not achieve the same spatial release as the normal-hearing listeners at equivalent sensation levels. In addition, the amount of informational masking created in each listener group differed, confounding the interpretation of the results. However, because listeners with hearing loss were able to benefit significantly from the perceptual effect of spatial separation, it is important to design and fit auditory prostheses in a way that retains the cues that underlie this perceptual effect. The most basic approach would be to provide bilateral amplification for bilateral hearing losses, when appropriate.

ACKNOWLEDGMENTS

This work was supported by Grant Nos. DC00100, DC04545, and DC04663 from NIH/NIDCD and by the Boston University Hearing Research Center. The authors are grateful to Andrew Brughera for critical technical assistance and to Robert Shannon and Qian-Jie Fu of House Ear Institute for providing the cochlear implant simulation program. The authors also thank Armin Kohlrausch and two anonymous reviewers for helpful comments on an earlier version of this paper.

¹In the current document, the general term “hearing impaired,” “hearing impairment,” or “hearing loss” will always refer to sensorineural hearing loss, presumably of cochlear origin.

²Although unintelligible, we cannot rule out the possibility that the DBN masker produced a small amount of informational masking because of the trial-to-trial uncertainty about the frequencies of the bands comprising the masker.

³The temporal modulation characteristics of DBS and DBN maskers have recently been studied by Kidd *et al.* (2004). They found that the envelope spectra of the two are similar with maxima in the 1–5-Hz region and a reduction in magnitude above 10 Hz. The envelope spectrum for the DBN has a slightly higher dc component than the DBS and varies smoothly over the range measured. The DBS masker, however, has small peaks roughly corresponding to modulation frequencies that are the inverse of the durations expected of syllables and words. The measured correlation between the envelopes of the DBS and the DBN maskers derived from the same speech sample was near zero.

⁴This statement is made based on S/Ms at the midpoints of the psychometric functions for the DBN masker played at 0° (–25 dB for the NH group; –12 dB for the HI group). Another way to compare groups in terms of energetic masking is to calculate the threshold shift from quiet for the DBN masker. Using this method, the amount of energetic masking is still significantly larger in the HI group (20.7 dB) than the NH group (13.2 dB), but the difference between groups is smaller (7 dB) than when calculated using S/M (13 dB).

Alexander, J. M., and Lutfi, R. A. (2003). “Informational masking for constant SL and SPL maskers in normal-hearing and hearing-impaired listeners,” presented at the 26th Annual Mid-Winter Meeting of the Association for Research in Otolaryngology, Daytona Beach, FL.

Arbogast, T. L., Mason, C. R., and Kidd, Jr., G. (2002). “The effect of spatial separation on informational and energetic masking of speech,” *J. Acoust. Soc. Am.* **112**, 2086–2098.

Bolia, R. S., Nelson, W. T., Ericson, M. A., and Simpson, B. D. (2000). “A speech corpus for multitalker communications research,” *J. Acoust. Soc. Am.* **107**, 1065–1066.

Bronkhorst, A. W., and Plomp, R. (1988). “The effect of head-induced interaural time and level differences on speech intelligibility in noise,” *J. Acoust. Soc. Am.* **83**, 1508–1516.

Bronkhorst, A. W., and Plomp, R. (1989). “Binaural speech intelligibility in noise for hearing-impaired listeners,” *J. Acoust. Soc. Am.* **86**, 1374–1383.

Bronkhorst, A. W., and Plomp, R. (1990). “A clinical test for the assessment of binaural speech perception in noise,” *Audiology* **29**, 275–285.

Bronkhorst, A. W., and Plomp, R. (1992). “Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing,” *J. Acoust. Soc. Am.* **92**, 3132–3139.

Brungart, D. S. (2001). “Informational and energetic masking effects in the perception of two simultaneous talkers,” *J. Acoust. Soc. Am.* **109**, 1101–1109.

Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. (2001). “Informational and energetic masking effects in the perception of multiple simultaneous talkers,” *J. Acoust. Soc. Am.* **110**, 2527–2538.

Brungart, D. S., Simpson, B. D., Darwin, C. J., Arbogast, T. L., and Kidd, Jr., G. (2005). “Across-ear interference from parametrically-degraded synthetic speech signals in a dichotic cocktail-party listening task,” *J. Acoust. Soc. Am.* **117**, 292–304.

Carhart, R., Tillman, T. W., and Greetis, E. S. (1968). “Perceptual masking in multiple sound backgrounds,” *J. Acoust. Soc. Am.* **45**, 694–703.

Cherry, E. C. (1953). “Some experiments on the recognition of speech with one and two ears,” *J. Acoust. Soc. Am.* **25**, 975–979.

Duquesnoy, A. J. (1983). “Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons,” *J. Acoust. Soc. Am.* **74**, 739–743.

Durlach, N. I., Mason, C. R., Kidd, Jr., G., Arbogast, T. L., Colburn, H. S., and Shinn-Cunningham, B. G. (2003a). “Note on informational masking,” *J. Acoust. Soc. Am.* **113**, 2984–2987.

Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., and Kidd, Jr., G. (2003b). “Informational masking: counteracting the effects of stimulus uncertainty by decreasing target-masker similarity,” *J. Acoust. Soc. Am.* **114**, 368–379.

Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (2001). “Spatial release from informational masking in speech recognition,” *J. Acoust. Soc. Am.* **109**, 2112–2122.

Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (1999). “The role of perceived spatial separation in the unmasking of speech,” *J. Acoust. Soc. Am.* **106**, 3578–3588.

Glasberg, B. R., and Moore, B. C. J. (1986). “Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments,” *J. Acoust. Soc. Am.* **79**, 1020–1033.

Grose, J. H., and Hall, III, J. W. (1996). “Perceptual organization of sequential stimuli in listeners with cochlear hearing loss,” *J. Speech Hear. Res.* **39**, 1149–1158.

Hawley, M. L. (2000). “Speech intelligibility, localization and binaural hearing in listeners with normal and impaired hearing,” Ph.D. dissertation, Biomedical Engineering, Boston University, Boston, MA.

Hawley, M. L., Litovsky, R. Y., and Culling, J. F. (2004). “The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer,” *J. Acoust. Soc. Am.* **115**, 833–843.

Kidd, Jr., G., Mason, C. R., and Arbogast, T. L. (2002b). “Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns,” *J. Acoust. Soc. Am.* **111**, 1367–1376.

Kidd, Jr., G., Mason, C. R., and Gallun, F. J. (2005). “Combining energetic and informational masking for speech identification,” *J. Acoust. Soc. Am.* (under review).

Kidd, Jr., G., Arbogast, T. L., Mason, C. R., and Walsh, M. (2002a). “Informational masking in listeners with sensorineural hearing loss,” *J. Assoc. Res. Otolaryngol.* **3**, 107–119.

Kidd, Jr., G., Mason, C. R., Brughera, A., and Chiu, C.-Y. P. (2003). “Discriminating harmonicity,” *J. Acoust. Soc. Am.* **114**, 967–977.

Kidd, Jr., G., Mason, C. R., Rohtla, T. L., and Deliwala, P. S. (1998). “Release from masking due to spatial separation of sources in the identification of non-speech auditory patterns,” *J. Acoust. Soc. Am.* **104**, 422–431.

Leek, M. R., and Summers, V. (1993). “Auditory filter shapes of normal-hearing and hearing-impaired listeners in continuous broadband noise,” *J. Acoust. Soc. Am.* **94**, 3127–3137.

Levitt, H. (1971). “Transformed up-down methods in psychoacoustics,” *J. Acoust. Soc. Am.* **49**, 467–477.

Neff, D. L., and Green, D. M. (1987). “Masking produced by spectral uncertainty with multi-component maskers,” *Percept. Psychophys.* **41**, 409–415.

- Noble, W., Byrne, D., and Ter-Horst, K. (1997). "Auditory localization, detection of spatial separateness, and speech hearing in noise by hearing impaired listeners," *J. Acoust. Soc. Am.* **102**, 2343–2352.
- Oh, E., and Lutfi, R. A. (1998). "Nonmonotonicity of informational masking," *J. Acoust. Soc. Am.* **104**, 3489–3499.
- Peissig, J., and Kollmeier, B. (1997). "Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners," *J. Acoust. Soc. Am.* **101**, 1660–1670.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**, 303–304.
- Watson, C. S. (1987). "Uncertainty, informational masking and the capacity of immediate auditory memory," in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Lawrence Erlbaum, Hillsdale, NJ), pp. 267–277.
- Zurek, P. M. (1993). "Binaural advantages and directional effects in speech intelligibility," in *Acoustical Factors Affecting Hearing Aid Performance*, edited by G. A. Studebaker and I. Hochberg (Allyn and Bacon, Boston), pp. 255–276.
- Zwicker, E., and Henning, G. B. (1984). "Binaural masking-level differences with tones masked by noises of various bandwidths and levels," *Hear. Res.* **14**, 179–183.

A Speech Intelligibility Index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners

Koenraad S. Rhebergen^{a)}

Department of Clinical and Experimental Audiology, Academic Medical Center, Room D2-223, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands

Niek J. Versfeld^{b)}

Department of Clinical and Experimental Audiology, Academic Medical Center, Room D2-330, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands

(Received 1 March 2004; revised 27 December 2004; accepted 31 December 2004)

The SII model in its present form (ANSI S3.5-1997, American National Standards Institute, New York) can accurately describe intelligibility for speech in stationary noise but fails to do so for nonstationary noise maskers. Here, an extension to the SII model is proposed with the aim to predict the speech intelligibility in both stationary and fluctuating noise. The basic principle of the present approach is that both speech and noise signal are partitioned into small time frames. Within each time frame the conventional SII is determined, yielding the speech information available to the listener at that time frame. Next, the SII values of these time frames are averaged, resulting in the SII for that particular condition. Using speech reception threshold (SRT) data from the literature, the extension to the present SII model can give a good account for SRTs in stationary noise, fluctuating speech noise, interrupted noise, and multiple-talker noise. The predictions for sinusoidally intensity modulated (SIM) noise and real speech or speech-like maskers are better than with the original SII model, but are still not accurate. For the latter type of maskers, informational masking may play a role. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1861713]

PACS numbers: 43.71.An, 43.66.Ba, 43.71.Gv, 43.72.Kb [PFA]

Pages: 2181–2192

I. INTRODUCTION

In daily life, speech is not always equally intelligible due to the presence of background noise. This noise may mask part of the speech signal such that not all speech information is available to the listener. In order to be able to predict the speech intelligibility under such masking conditions, French and Steinberg (1947), Fletcher and Galt (1950), and later Kryter (1962a, b) initiated a calculation scheme, known as the Articulation Index (AI), which at present still is used by a number of investigators (Rankovic, 1998, 2002; Hogan and Turner, 1998; Müsch and Buus, 2001; Brungart, 2001; Turner and Henry, 2002; Dubno *et al.*, 2002, 2003). In 1984, Pavlovic and others (Dirks *et al.*, 1986; Kamm *et al.*, 1985; Pavlovic, 1984, 1987; Pavlovic and Studebaker, 1984; Pavlovic *et al.*, 1986; Studebaker *et al.*, 1987, 1994) started to re-examine the AI calculation scheme, which has led to a new method accepted as the ANSI S3.5-1997 (1997). Since its revision in 1997, the method is named the Speech Intelligibility Index (SII).

For a given speech-in-noise condition, the SII is calculated from the speech spectrum, the noise spectrum, and the listener's hearing threshold. Both speech and noise signal are filtered into frequency bands. Within each frequency band the factor audibility is derived from the signal-to-noise ratio (SNR) in that band indicating the degree to which the speech is audible. Since not all frequency bands contain an equal

amount of speech information (i.e., are not equally important for intelligibility), bands are weighted by the so-called band-importance function. The band-importance function indicates to which degree each frequency band contributes to intelligibility. It depends on the type of speech material involved (e.g., single words or sentences), and other factors. Finally, the SII is determined by accumulation of the audibility across the different frequency bands, weighted by the band-importance function. The resulting SII is a number between zero and unity. The SII can be seen as the proportion of the total speech information available to the listener. An SII of zero indicates that no speech information is available to the listener, an SII of unity indicates that all speech information is available. Model parameters have been chosen such that the SII is highly correlated to intelligibility. The SII model has been developed to predict the *average* speech intelligibility for a given speech-in-noise condition; it does not attempt to predict the intelligibility of the individual utterances (phonemes or words) of a speech fragment. Also, speech redundancy or contextual effects, which are inherent to meaningful speech, are captured in the SII model by choice of the model parameters. Higher speech redundancy simply results in less information (i.e., a lower value for the SII) required for understanding the speech message. Within the context of the present paper, an important observation is that the existing SII model does not take into account any fluctuation in the masking noise, since the SII is computed from the long-term speech and noise spectrum. Therefore, the SII

^{a)}Electronic mail: k.s.rhebergen@amc.uva.nl

^{b)}Electronic mail: n.j.versfeld@amc.uva.nl

is independent of the amount of fluctuations in the noise signal.

Numerous papers have reported on experiments dealing with speech intelligibility in fluctuating noise. In almost all cases, normal-hearing listeners perform better in conditions with fluctuating noise compared to those with stationary noise of the same rms level (Miller, 1947; Miller and Licklider, 1950; Licklider and Guttman, 1957; de Laat and Plomp, 1983; Duquesnoy, 1983; Festen, 1987, 1993; Festen and Plomp, 1990; Gustafsson and Arlinger, 1994; Bacon *et al.*, 1998; Peters *et al.*, 1998; Brungart, 2001; Versfeld and Dreschler, 2002; Dubno *et al.*, 2002; Nelson *et al.*, 2003). In many cases, this finding has been phenomenologically explained by stating that the listener is “able to catch glimpses of the speech during the short silent periods of the masking noise” (Howard-Jones and Rosen, 1992, 1993; Festen, 1993; Peters *et al.*, 1998). Recently, Oxenham and co-workers (Oxenham and Plack, 1997; Plack and Oxenham, 1998; Oxenham *et al.*, 2004) proposed that the nonlinear behavior of the basilar membrane enables increased gain during the silent periods, allowing increased audibility. In hearing-impaired subjects, this nonlinear behavior is less or even absent, which results in decreased audibility during absence of masking noise.

So far, the SII model has been validated only for stationary masking noises, for which it works well. However, it fails to predict speech intelligibility accurately in the case of fluctuating noise maskers (Festen and Plomp, 1990; Houtgast *et al.*, 1992; Versfeld and Dreschler, 2002). Other methods, such as the Speech Transmission Index (STI, Steeneken and Houtgast, 1980), or even the speech-based STI (van Wijn-gaarden, 2002) also fail at this point. To our knowledge, there is still no method that can predict the speech intelligibility in fluctuating noise accurately. Yet, since most real-life noises do exhibit strong variations over time, there is great interest in a procedure that is able to predict speech intelligibility in fluctuating noises adequately.

In the present paper, an extension to the SII model is proposed in order to be able to predict the speech intelligibility not only in stationary noise, but also in fluctuating noise. The extension consists of an approach where, for a given condition, both speech and noise signal are partitioned into small time frames. Within each time frame, the conventional SII is determined, yielding the speech information available to the listener at that time frame. Next, the SII values of these time frames are averaged, resulting in the SII for that particular noise type. It is hypothesized that this averaged SII is closely related to the speech intelligibility for that condition.

In the next section, an outline of the existing SII model is given. It is followed by a detailed description of the extensions to the existing model, which are introduced to allow predictions of the speech intelligibility in fluctuating noises as well. In extending the SII model, attention has been given to stay as close as possible to the original SII model, thus making as few adaptations as possible. In the choice of the model parameters, this paper concentrates on experiments where speech intelligibility has been assessed with the method of the so-called speech reception threshold (SRT), as

described by Plomp and Mimpen (1979). With this method, short everyday sentences are used as speech materials. In Sec. II C the SRT method is described in some detail. Next (in Sec. III) data from the literature are used to evaluate the extended SII model. Finally, in Sec. IV, predictions and limitations of the extended SII model will be discussed.

II. MODEL DESCRIPTION

A. The SII model

A detailed description of the SII model is given in ANSI S3.5-1997 (1997). Here, a brief overview is given so that in the next section the extensions to the existing model are easier to follow.

The SII model basically calculates the average amount of speech information available to a listener. To that extent, the model uses the long-term averaged speech spectrum and the long-term averaged noise spectrum as input. Both speech and noise spectrum are defined as the spectrum level (in dB/Hz) at the eardrum of the listener. Within the model, an option exists to partition the speech and noise spectrum into octave bands, one-third-octave bands, or critical bands. In this paper, spectra are partitioned into critical bands (given in Table I of the ANSI S3.5-1997 standard), although the other two options are equally valid. Within each critical band, the spectrum level is separately determined for both speech and noise. Next, correction factors are taken into account for effects such as upward spread of masking for both speech and noise, inaudibility due to the auditory threshold for pure tones, and distortion due to excessive high speech or noise levels. Then, within each frequency band, the difference between the speech and noise level (signal-to-noise ratio, or SNR) is calculated and this value is multiplied with the so-called band-importance function, which results in the proportion of information in that band that is available to the listener. The band-importance function may depend on the type of speech materials (e.g., sentences or words), or level. Finally, these values are added, yielding the Speech Intelligibility Index (SII), or the amount of speech information available to the listener. For normal-hearing listeners, the SII has proven to be closely related to the average intelligibility in a given condition where speech is masked by a stationary noise masker (Pavlovic, 1987).

B. Extension to the SII model

Since the SII model uses the long-term averaged speech and noise spectrum as input, all temporal characteristics of these signals are lost. As mentioned in the Introduction, large differences in intelligibility exist between masking noises that differ from each other solely with respect to temporal fluctuations (e.g., steady-state versus fluctuating noise). In this section, an extension is presented that does take the temporal characteristics of the masking noise into account. In essence, the SII model is adapted such that the SII is calculated within small time frames, after which the average SII is calculated.

A block diagram of the calculation scheme is presented in Fig. 1. Both speech and noise are analyzed separately for the SII calculation. Although, in principle, regular speech could be used as the speech input signal, speech-shaped

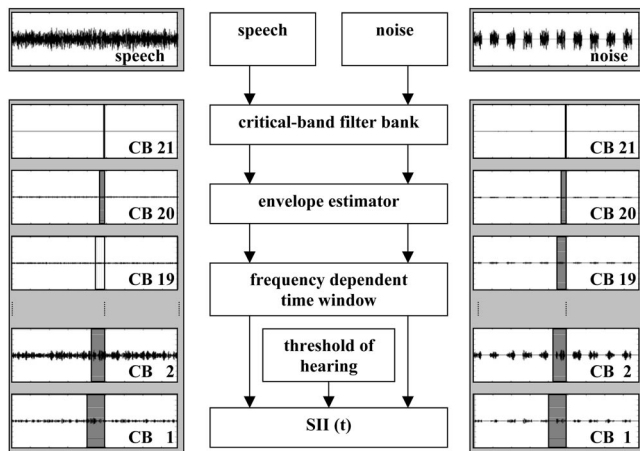


FIG. 1. Schematic overview of the calculation scheme for the extended SII model. A detailed description is given in the main text. The input speech signal (stationary Gaussian noise with the long-term average spectrum of speech) and input noise (in this example interrupted noise with the long-term average spectrum of speech) are separately filtered by a 21 critical-band (CB) filter bank. The envelope of the input speech and noise are estimated in every CB (1–21); the instantaneous intensity is estimated in a frequency-dependent time window, as indicated by the shaded bars (CB1=35 ms to CB21=9.4 ms). Every 9.4 ms an SII is calculated as described by ANSI S3.5-1997. For each of the approximately 200 steps (of 9.4 ms), the instantaneous SII(t) is determined (sentence of about 2 s). Last, the SII for that speech-in-noise condition is determined by averaging across all instantaneous SII(t) values.

noise (i.e., stationary Gaussian noise with the long-term average spectrum of speech) was used. The main reason for this is that, in combination with stationary noise as a noise masker, all SII values are identical to those obtained with the existing SII model. This prerequisite is not easily fulfilled when normal speech signals would be used.

The SII is in principle designed to predict the average intelligibility of speech in noise and not the intelligibility of individual words or phonemes. In any case, the SII is badly defined in case of silent periods occurring within the normal speech signal because, regardless of the masking noise, the SII will always be zero. Thus, even when a speech signal is presented at a clear level without any masking noise, the SII based on regular speech never will reach unity, due to the inherent silent periods in the speech signal. Moreover, problems will occur if one considers the silent periods between sentences. It is clear that large differences in SII may occur when the silent periods between sentences vary, whereas the actual intelligibility should not be different.

The most straightforward approach to determine the SII within small time frames is to window the speech and noise signal at a given point in time, calculate the frequency spectrum (by means of a fast Fourier transform, FFT), and derive an SII from the resulting speech and noise spectrum and the threshold of hearing. However, in order to be able to track the perceptually relevant fluctuations over time, the window length should be small enough. This means that the time window should have a duration of several milliseconds, which is the temporal resolution for normal-hearing listeners based on gap-detection thresholds in the higher frequency bands (Plomp, 1964; Shailer and Moore, 1983, 1987; Glasberg and Moore, 1992; Eddins *et al.*, 1992; Oxenham and

Moore, 1994, 1997; Moore *et al.*, 1996; Plack and Oxenham, 1998; Moore, 1997). Unfortunately, such a short time window leads to the signal-analytical problem that the level in the lower frequency bands is not estimated accurately. On the other hand, a longer time window leads to a poorer grasp of the temporal variations of the signal.

It is known that the temporal resolution of the auditory system is frequency dependent (Shailer and Moore, 1983, 1987). Time constants (i.e., integration times) for the lower frequency bands are larger than those for the higher bands. To overcome the analysis problems on the one hand, and to stay close to the characteristics of the auditory system with respect to temporal resolution on the other hand, the signal was first filtered into 21 critical bands, and the window length was chosen to be relatively short in the higher bands and relatively long in the lower bands. Since in the original SII calculations the frequency bands are essentially nonoverlapping (after all, the intensity within each filter band was derived from the frequency spectrum), a FIR filter bank of order 200 [MATLAB function `firl(200,Wn)`] was used to filter the entire speech and noise signal into the separate bands. Within each band, the temporal envelope was determined by means of a Hilbert transform. At a given time frame, rectangular windows were used with window lengths ranging from 35 ms at the lowest band (150 Hz), to 9.4 ms at the highest band (8000 Hz). These window lengths were taken from Moore (1997, Chap. 4) for gap detection and have been multiplied by 2.5. The factor 2.5 was chosen to provide a good fit to the present data set, as will be discussed below. The windows were aligned such that they ended simultaneously. Within each time frame the intensity was determined, and these, together with the absolute threshold for hearing were used as input to calculate the instantaneous SII, for that given time frame. To calculate the SII, the so-called speech perception in noise (SPIN) weighting function (ANSI S3.5-1997, 1997, Table B.1) was used. This choice seems to be valid, since the speech materials of Plomp and Mimpen (1979) are closely related to the SPIN materials with respect to sentence length and redundancy. Last, the SII for the speech-in-noise condition under consideration was determined by averaging across all instantaneous SII values.

C. Speech reception threshold

In the present paper, the proposed extension to the SII model was evaluated using existing data from the literature. The data differ from each other with respect to a number of variables that all can have an effect on intelligibility, hence on the parameter settings of the SII model. For example, it is known that the type of speech material (monosyllables, words, sentences, etc.), open or closed response set, and native or non-native language acquisition can have a large effect on intelligibility (Bosman and Smoorenburg, 1995; Drullman and Bronkhorst, 2000; van Wijngaarden, 2003). Next, similarity between masker and target, e.g., in the case where both target and masker consist of a male voice (Bronkhorst and Plomp, 1992; Bronkhorst, 2000), has a detrimental effect on the actual threshold (i.e., the signal-to-noise ratio that results in just-intelligible speech). Also, the experimental paradigm influences threshold to a large extent.

The adaptive SRT procedure according to Plomp and Mimpen (1979), and the Just to Follow Conversation (Hygge *et al.*, 1992; Larsby and Arlinger, 1994) result in different threshold levels for the same speech material. Additionally, differences in data acquisition (e.g., strictness of sentence scoring) may have an effect on threshold level. Furthermore, different presentation methods (through headphones, loudspeakers, monaural, binaural, diotic, or dichotic presentation) evidently affect threshold level. If one considers masking noises bearing silent periods, it is likely that, even within a group of normal-hearing subjects, differences in hearing level may affect audibility, and thus intelligibility. Finally, when dealing with spectral differences between masker and target, the method used for calibrating signal levels (e.g., rms, dBA) may have a clear effect.

To enable a comparison between data obtained in different studies, in the present study only thresholds are used that were obtained with the so-called speech reception threshold (SRT) method for sentences, as described by Plomp and Mimpen (1979). Speech materials consist of simple everyday sentences, having a length of 8 to 9 syllables (Plomp and Mimpen, 1979; Nilsson *et al.*, 1994; Versfeld *et al.*, 2000). The SRT is defined as the signal-to-noise ratio (SNR) needed for 50% sentence intelligibility. The SRT is estimated as described by Plomp and Mimpen (1979): A list of 13 sentences, unknown to the listener, is monaurally presented via headphones. The masking noise is presented at a fixed level, whereas the sentence level is varied adaptively. The first sentence starts at a very unfavorable SNR, and is repeated each time at a 4-dB higher level until the listener is able to repeat every word of this sentence exactly. The SNR of the 12 remaining sentences is varied adaptively with a step size of 2 dB using a one-up, one-down procedure. The SNR of the next sentence is increased by 2 dB after an incorrect response and decreased by 2 dB after a correct response. The average adjusted SNR of sentence 4 through 13 is adopted as the SRT for that particular noise condition. With the speech material of Plomp and Mimpen (1979), normal-hearing listeners require an SNR in stationary speech-shaped noise of -5 to -4 dB, which corresponds to an SII between 0.3 and 0.4 (Steeneken, 1992; Bronkhorst, 2000; Noordhoek, 2000; Versfeld and Dreschler, 2002; van Wijngaarden, 2002, 2003). This means that roughly one-third of the speech information is required to the normal-hearing listener (i.e., the SII is between 0.3 and 0.4) to reach the SRT for these sentences.

III. MODEL PREDICTIONS

A. Steady-state speech noise

Speech intelligibility in stationary speech-shaped noise can be well predicted by the existing SII model. There are numerous papers dealing with the SRT in stationary speech noise, and all report for normal-hearing listeners at a fixed noise level between 60 and 80 dBA an SRT for sentences of approximately -4.5 dB (de Laat and Plomp, 1983; Middelweerd *et al.*, 1990; Festen, 1987; Festen and Plomp, 1990; ter Keurs *et al.*, 1993; Versfeld and Dreschler, 2002; Neij-

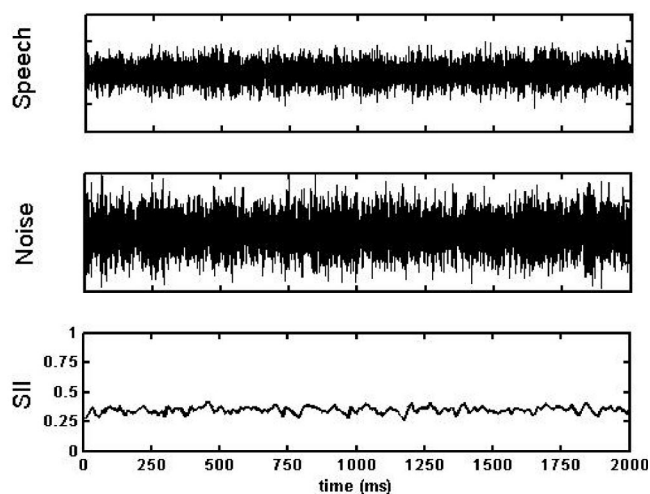


FIG. 2. Representation of the SII with the extended SII model for a speech-in-noise sample of 2 s. The upper panel represents a speech signal of a female speaker. The middle panel represents a stationary speech-shaped masking speech noise. The noise has been scaled to 60 dBA. The target has been scaled to 55.5 dBA, which results in an SNR of -4.5 dB. The lower panel displays the resulting instantaneous SII as a function of time. The SII averaged across time is equal to 0.35.

huis, 2002). For speech in stationary speech noise, an SRT of -4.5 dB results for the existing SII model in an SII value of 0.35.

Figure 2 displays the results of a calculation with the extended SII model for speech in stationary speech noise. The upper panel in Fig. 2 displays the waveform of a speech signal representation (that is—a stationary speech-shaped noise signal instead of an actual speech signal, as discussed in the previous section) with a duration of 2 seconds, presented at a level of 55.5 dBA. Here, speech noise was taken from Versfeld *et al.* (2000) for the female speaker. The middle panel shows a 2-s sample of the stationary speech-shaped noise masker derived from the same female speaker, at a level of 60 dBA. The lower panel in Fig. 2 shows the resulting instantaneous SII, where the SII has been determined every 9.4 ms. Due to the fact that speech and noise signal are uncorrelated (different noise samples), small fluctuations in the instantaneous SII occur. It is easy to see that the SII, averaged across the 2-s sample, is between 0.3 and 0.4. In fact, the average is 0.35, which is identical to the value obtained by the existing SII model. Many conditions with speech in stationary noise have been studied, and all calculations show that neither speech type nor noise type result in differences between the existing SII model and the present extended SII model. In conclusion, the extended SII model yields exactly the same results as the existing SII model, as long as a stationary masking noise is used.

B. Speech noise with a speech-like modulation spectrum

As discussed above, the existing SII model is not able to correctly predict intelligibility for speech in modulated noise. This section deals with speech intelligibility for speech in noise with a speech-like spectrum and a single-speaker modulation spectrum. The generation of this type of noise is described by Festen and Plomp (1990). With normal-hearing

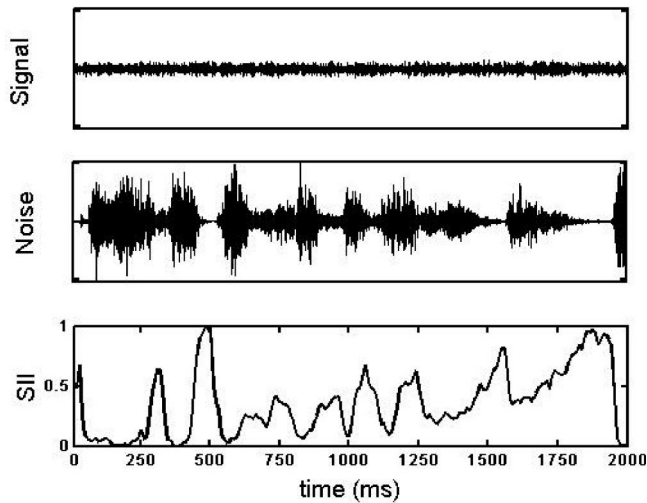


FIG. 3. Representation of the SII with the extended SII model for a speech-in-noise sample of 2 s. The upper panel represents a speech signal of a female speaker. The middle panel represents a fluctuating speech-shaped masking speech noise, as used by Festen and Plomp (1990). The noise has been scaled to 60 dBA. The target has been scaled to 48 dBA, which results in an SNR of -12 dB. The lower panel displays the resulting instantaneous SII as a function of time. The SII averaged across time is equal to 0.35.

subjects, several papers report for this condition an SRT around -12 dB (Festen and Plomp, 1990; ter Keurs *et al.*, 1993; Versfeld and Dreschler, 2002; Neijenhuis *et al.*, 2002), when the noise level is between 60 and 80 dBA. Computations with the existing SII model yield a score of 0.089, which is far too low. Figure 3 displays the results of the calculations with the extended SII model, similar to the previous section. The upper panel displays the waveform of a speech signal (again, taken as a stationary speech-shaped noise signal) with a duration of 2 seconds, presented at a level of 48 dBA. The middle panel shows a 2-s sample of the modulated speech noise masker, at a level of 60 dBA. The lower panel in Fig. 3 shows the resulting instantaneous SII, where, in contrast to the findings in Fig. 2, the SII value greatly varies over time. It ranges from values close to zero (at points in time where the speech is entirely masked by the masking noise) to values near unity (at points where the masking noise is momentarily absent). The lower panel thus denotes the amount of speech information available to the listener as a function of time. Averaging across time results in an SII score of 0.35. Because large fluctuations exist over time, a suitably long period has to be chosen to average across. The time interval required to reach stable values for the SII depends on the periodicity, or alternatively, randomness, of the signal as well as on the modulation frequencies in the masking signal. With the present type of masking noise, where the modulations are most prominent near 4 Hz, a period of 2 s appears to be long enough to reach a between-samples standard deviation for the SII of 0.0056. Increasing the period to 4 s decreases the standard deviation of the SII to 0.0030.

Figure 4 displays the SII as a function of the SNR. Here, the masking noise has been kept fixed at 60 dBA, and the level of the speech has been varied between 30 and 80 dBA (thus between SNRs of -30 and $+20$ dB). With stationary speech noise (denoted as filled symbols in Fig. 4) the SII

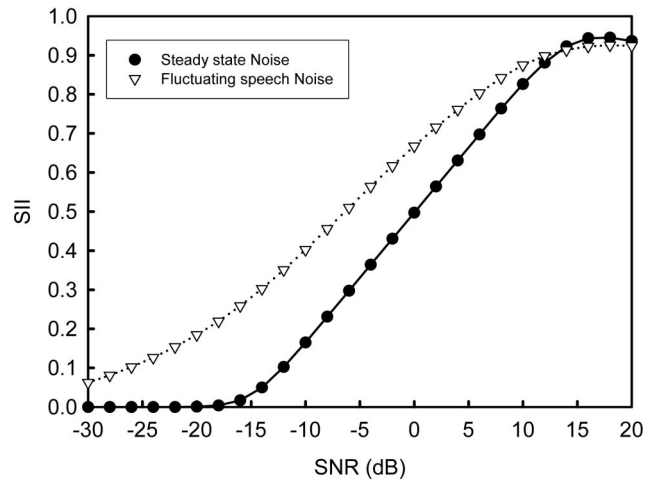


FIG. 4. SII as a function of SNR as calculated with the extended SII model. Filled symbols denote calculations with a stationary noise masker with the long-term spectrum of the female target speaker. Open symbols denote calculations with a fluctuating noise masker with the long-term spectrum of the female target speaker and a speech-like modulation spectrum. The level of the noises was set to 60 dBA.

starts to deviate from zero as the SNR reaches a value of -15 dB and increases almost linearly with the SNR up to a value of $+15$ dB. At this value, the speech level is about 75 dBA, and the distortion factor in the SII model prevents the SII from reaching unity. The behavior of the SII as a function of SNR with stationary noise is identical for the existing and the extended SII model. Differences between the two models arise when fluctuating noise is used as a masker. Since the existing SII model does not take the amplitude modulations in the noise masker into account, the SII as calculated with the existing SII model will be identical to that calculated for stationary noise. The SII as a function of SNR for fluctuating noise predicted by the extended SII model is given with open symbols in Fig. 4. Even at very low signal-to-noise ratios, there is still some speech information available to the listener and the SII exceeds zero. Increasing the SNR causes the SII to increase, but the slope of the function is not as steep as that calculated for speech in stationary noise. Again, at higher speech levels, the distortion factor of the SII model causes the function to level off, such that the SII does not reach unity. An important observation seen in Fig. 4 is that a constant SII value of 0.35 (the information required to reach threshold) results in an SRT of -4.5 dB for stationary masking noise and -12 dB for fluctuating masking noise.

C. Interrupted speech noise

de Laat and Plomp (1983) measured SRTs for sentences in interrupted (gated) speech noise with a duty cycle of 50%. Modulation frequency was 10 Hz. Masking noise was presented at 65, 75, or 85 dBA. Figure 5 displays the calculations with the extended SII model, similar to Figs. 2 and 3. The upper and middle panel show the speech signal and masking noise signal, respectively. Signal and noise level are 42 and 65 dBA, respectively. The SNR thus is -23 dB. The lower panel shows the SII as a function of time. As seen earlier, the SII is close to zero when the masking noise is present, and is close to unity when the masking noise is absent. Due to the longer integration times in the lower fre-

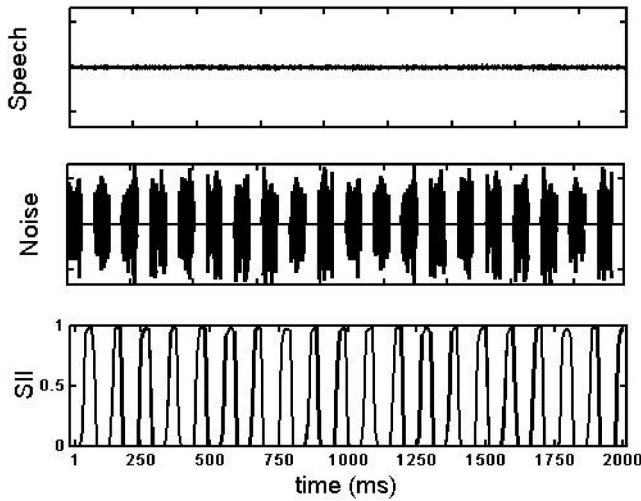


FIG. 5. Representation of the SII with the extended SII model for a speech-in-noise sample of 2 s. The upper panel represents a speech signal of a female speaker. The middle panel represents an interrupted speech-shaped masking speech noise, as used by de Laat and Plomp (1983). The noise has been scaled to 65 dBA. The target has been scaled to 42 dBA, which results in an SNR of -23 dB. The lower panel displays the resulting instantaneous SII as a function of time. The SII averaged across time is equal to 0.35.

quency bands, the SII does not change as rapidly as the interrupted noise, but rather smears out over time. Again, the SII averaged across time is equal to 0.35.

Figure 6 displays the SII as a function of SNR for stationary speech noise (filled symbols), and for the three conditions with 10-Hz interrupted noise used in de Laat and Plomp (1983, open symbols; noise at 65, 75, and 85 dBA). At low SNRs (between -15 and -35 dB), speech is entirely masked at moments when the masking noise is present, and it is audible in the gaps. Due to the gaps in the masking noise, values for the SII are relatively independent of SNR and are still quite large, on the order of 0.3. At even lower SNRs (below -35 dB), SII eventually decreases to zero, due to the fact that the speech signal will fall below the absolute

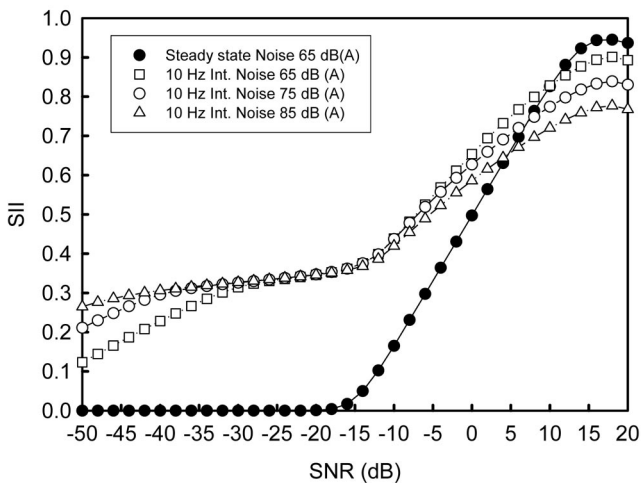


FIG. 6. SII as a function of SNR as calculated with the extended SII model. Filled symbols denote calculations with a stationary noise masker with the long-term spectrum of the female target speaker at a level of 60 dBA. Open squares, circles, and triangles denote calculations with the interrupted noise masker with the long-term spectrum of the female target speaker where the level of the noise was set to 65, 75, and 85 dBA, respectively.

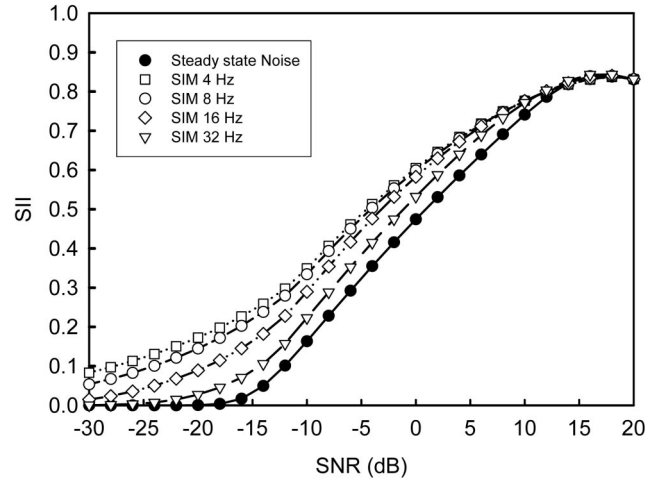


FIG. 7. SII as a function of SNR as calculated with the extended SII model. Filled symbols denote calculations with a stationary noise masker with the long-term spectrum of the female target speaker at a level of 75 dBA. Open squares, circles, diamonds, and triangles denote calculations with SIM noise as a masker at a level of 75 dBA, and a modulation frequency of 4, 8, 16, and 32 Hz, respectively.

threshold. Absolute threshold here has been taken equal to 0 dB (HL). At an SNR of -15 and larger, portions of the speech signal start to exceed the noise signal, and SII increases. Again, at high speech levels, distortion occurs which causes the function to level off. de Laat and Plomp (1983) found an SRT of -23 , -26 , and -29 dB at a presentation level of the noise of 65, 75, and 85 dBA, respectively. Figure 6 shows that for these conditions a large variation in the SNR causes only a slight variation in the SII. At time frames where the noise signal is present, no speech information is available; but at time frames where the noise masker is absent, the amount of speech information available is determined by the degree of temporal resolution (i.e., forward and backward masking) as well as by the absolute threshold of hearing. Nevertheless, while computations with the existing SII model give an SII of zero, the extended SII model results in values near 0.35.

D. Sinusoidally intensity-modulated speech noise

Festen (1987) measured the SRT for sentences in 100% sinusoidally intensity-modulated (SIM) speech noise. At a presentation level of the noise of 75 dBA he found SRTs of -7.5 , -9 , -10 , -10.2 , and -4 dB for modulation frequencies of 4, 8, 16, 32, and “infinity” Hz (steady state), respectively. Figure 7 displays the SII as a function of SNR for stationary speech noise (filled symbols), and for four conditions with SIM noise used in the study of Festen (1987, open symbols). Computations with the extended SII model, given an SII of 0.35, result in SRTs of -10 , -9 , -8 , -6.3 , and -4 dB for the above-mentioned conditions. The predicted SRT in a 4-Hz SIM noise with the extended SII model seems to be lower compared to SRT values obtained by Festen (1987). Furthermore, the predicted SRT in a 16- or a 32-Hz SIM noise with the extended SII model seems to be higher compared to SRT values obtained by Festen (1987). Although the SRT values obtained with the extended SII model indicate an improvement over the existing model (which pre-

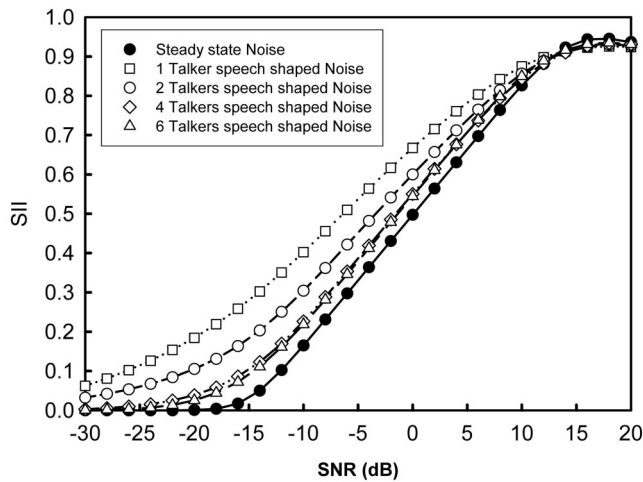


FIG. 8. SII as a function of SNR as calculated with the extended SII model. Filled symbols denote calculations with a stationary noise masker with the long-term spectrum of the female target speaker. Open squares, circles, diamonds, and triangles denote calculations with noise derived from a single, two, four, and six speakers speech-shaped noise. The level of the noises was set to 65 dBA.

dicts an SRT of -4 dB for all conditions), there are still some deviations. So far, no explanation can be given for this result.

E. Multiple-talker noise

There are numerous papers dealing with the SRT for speech in the presence of one or more competing talkers (e.g., Festen and Plomp, 1990; Bronkhorst and Plomp, 1992; Bronkhorst, 2000; Drullman and Bronkhorst, 2000; Brungart, 2001; Brungart *et al.*, 2001, 2002). It is generally observed that the SRT becomes worse as the number of competing voices increases (Miller, 1947; Carhart *et al.*, 1969; Bronkhorst and Plomp, 1992), eventually resulting in the SRT for stationary speech noise. Bronkhorst and Plomp (1992) measured the SRT for sentences masked by speech-shaped noise modulated by the envelope derived from one, two, four, or six interfering speakers. Observed SRTs were -9.7 , -9.9 , -7.2 , and -6.4 dB, respectively. The stimuli, i.e., speech and fluctuating speech noise, were recorded with a KEMAR manikin and presented monaurally to the subjects. Figure 8 displays for the four conditions of Bronkhorst and Plomp (1992) calculations of the extended SII model as a function of the signal-to-noise ratio where it was attempted to simulate Bronkhorst and Plomp's (1992) speech-shaped noises. It shows that at an SII value fixed at 0.35, the SRT increases from -12 dB (for a single interfering speech shaped noise) to -6 dB (for six interfering speech-shaped noises). Although the masking noises were regenerated, since the original masking noises of Bronkhorst and Plomp (1992) were not available, the trend is similar to that reported in the original study.

IV. DISCUSSION

Figure 9 displays the relationship between the observed SRT (i.e., as measured in actual experiments) and the SRT as predicted by the extended SII model for all conditions de-

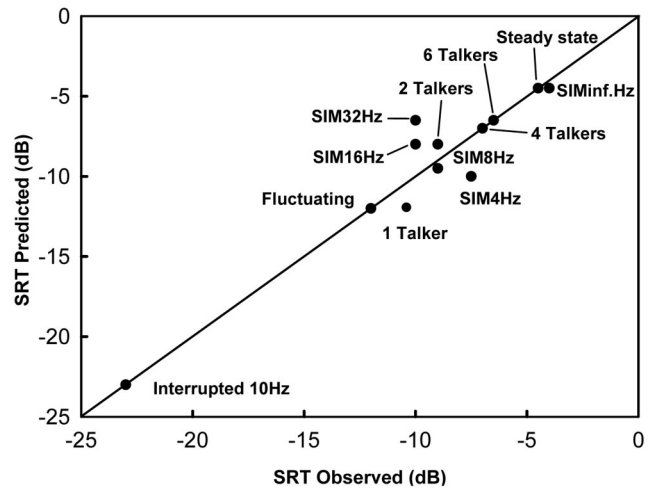


FIG. 9. For a number of different masking noises, the SRT (dB) predicted with the extended SII model is plotted as a function of the observed SRT (dB). Conditions are denoted in short in the figure.

scribed in the previous section, as well as some other conditions that will be discussed below. SRTs were calculated by taking the hearing loss fixed at 0 dB(HL) at all audiometric frequencies, and by setting the threshold value of the SII to 0.35. Different SRTs were obtained by taking the associated sample of the masking noise. The diagonal indicates the points where the observed and predicted SRT are equal. Points under the diagonal indicate an overestimation (with respect to performance) of the predicted SRT; points above the diagonal indicate that listeners generally perform better than predicted by the extended SII model. All predicted SRT values are within a few decibels of the diagonal, or even lie on the diagonal, indicating that the model does well with the present set of data. The extended SII model yields a substantial improvement over the existing model. Since the latter is insensitive to modulations in the masking noise, it thus predicts for practically all conditions an SRT of -4.5 dB. The most important finding of this paper is that average speech intelligibility in fluctuating noise can be modeled by averaging the amount of speech information across time.

If the data in Fig. 9 are considered in detail, some of the results obtained with the SIM noises of Festen (1987) seem to deviate to some degree from the diagonal. Festen (1987) found lowest SRTs for modulation frequencies of 16 and 32 Hz. His finding is in contrast with most data from the literature that indicate maximum performance at 10 Hz (Miller and Licklider, 1950; Licklider and Guttman, 1957; Gustafsson and Arlinger, 1994; Trine, 1995; Bronkhorst, 2000; Nelson *et al.*, 2003). The difference in the position of the minimum may be attributable to differences in stimulus type (gated noise versus SIM noise) and speech materials (word versus sentence scoring). There appears to be a large difference in the SRT results (about 16 dB) found by de Laat and Plomp (1983) and Festen (1987) obtained with about the same modulation frequencies [modulation frequency: 10 Hz; SRT: -26 dB for de Laat and Plomp (1983), compared to modulation frequency: 8 Hz: SRT -10 dB for Festen (1987)]. Festen (1987) suggested that this discrepancy can be due to the relatively broad and deep minimum in the inter-

rupted noise compared to that in the SIM noise (Fig. 2 from Festen, 1987). The SRT values, obtained with 16- and 32-Hz SIM noise are very similar, *viz.*, -10 dB, and are 2 to 3 dB better than predicted by the extended SII model. As for now, we have no explanation for this part of Festen's (1987) data. Increasing the modulation frequency of the SIM noise results in gaps that are sufficiently small such that they start to fall within the time window of the extended SII model (*i.e.*, smaller than 35 ms). This results in a decrease in performance, and finally performance will approach that of stationary noise. This condition is indicated by "SIMinf.Hz" in Fig. 9, and is close to the diagonal. Decreasing the modulation frequency to 8 Hz also results in a point close to the diagonal. However, a further decrease of the modulation frequency to 4 Hz again results in a deviation from the diagonal. The overestimation of the 4-Hz SIM noise may be accounted for by the fact that with these slow modulation rates, masking of complete words in a sentence can occur. This phenomenon has already been observed by Miller and Licklider (1950), who found optimal performance around modulation rates of 10 Hz. The mere fact that complete words are masked implies that the SRT procedure—where every word of the sentence needs to be repeated correctly—is unsuitable for these low modulation frequencies. Indeed, Trine (1995) shows that in the so-called Just-to-Follow-Conversation (JFC) procedure, the signal-to-noise ratio keeps on decreasing below modulation rates of 8 Hz. In this procedure, the subject is asked to adjust the level of speech in a fixed given noise masker such that he or she is able to "just follow" the speech. This procedure does not require the intelligibility of individual syllables, words, or even sentences. Therefore, the optimum performance for 8 Hz is a procedural artifact. Hence, to validate the extended SII model for masking noises comprising modulation rates of, say, 8 Hz and below, procedures other than the SRT procedure of Plomp and Mimpen (1979) should be utilized.

A. Effect of informational masking

The extended SII model may not be able to predict SRTs accurately in conditions where speech and masking noise interfere at a higher level. One example of such interference is when both target speech and masking noise are derived from the same speaker. In that condition, the listener is confused since he or she does not know which signal represents the target and which components of the signal represents the masker. Festen and Plomp (1990) describe a number of conditions where speech is masked by a single speaker or by multiple speakers. Indeed, performance for speech intelligibility in time-reversed masking speech is better than for forward-masking speech. This additional masking, on top of energetic masking, is called informational masking (Bronkhorst, 2000; Brungart, 2001; Brungart *et al.*, 2001): The spoken message of real interfering speech accounts for a rise in SRT.

In another experiment, Festen (1993) measured SRTs in other speech-like maskers. The target speech was uttered by a female speaker (of Plomp and Mimpen, 1979). The interfering speech consisted of comparable sentences from a male voice (Smooenburg, 1992). In the reference condition, the

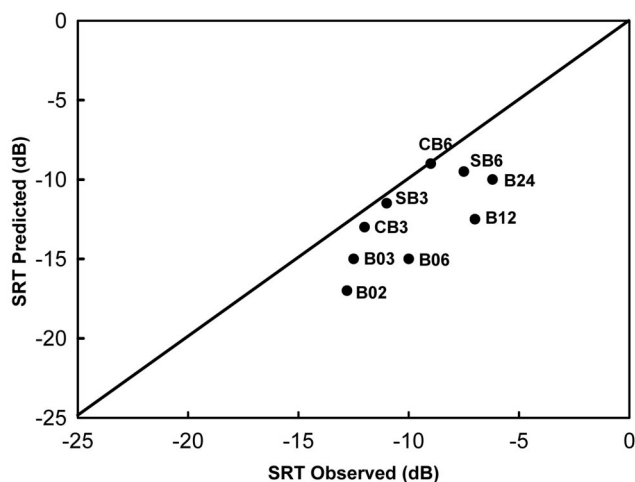


FIG. 10. The SRT (dB) predicted with the extended SII model is plotted as a function of the observed SRT (dB) for the noise maskers used in Festen (1993). Conditions are denoted by abbreviations in the figure. In conditions B02 through B24, conditions consisted of speech fragments that were manipulated by shifting individual frequency bands of the noise masker independently over time. In conditions CB3, CB6, SB3, and SB6, half of the speech masker was replaced by stationary speech noise. For further details the reader is referred to the main text.

interfering speech signal consisted of a concatenation of sentences, with no pauses between the sentences. Five other conditions were derived from this reference condition by first dividing the masking speech stream into 2, 3, 6, 12, or 24 separate frequency bands that next were independently shifted in time. One may see this masker as an addition of 2, 3, 4, 6, 12, or 24 speakers where the speech of the individual speakers does not overlap in frequency. The result is a masker that sounds very speech-like. The measured SRTs as well as the SRTs calculated with the extended SII model are displayed in Fig. 10. Different conditions are denoted as B02, B03, B06, B12, and B24, where the number denotes the number of frequency bands. The extended SII model appears to overestimate the observed SRT values of all conditions by 4 to 5 dB. Although speech and noise masker were well discernible, informational masking may have played a role, since the maskers still resembled running speech.

In addition to these conditions, Festen (1993) generated other maskers, where the upper 1/3 octave of each frequency band in the 3- and 6-band speech masker was replaced by noise of the same level as the time average of the original masker. Maskers therefore consisted half of stationary speech-shaped noise. The modulated part was either synchronous in time (labeled in Fig. 10 as "CB" for "constant bands") or shifted in time (labeled in Fig. 10 as "SB" for "shifted bands"). As can be seen in Fig. 10, the extended SII model is able to predict the SRT of all these noise conditions (CB3, CB6, SB3, and SB6) reasonably well, probably due to the fact that the masker is less speech-like.

In summary, when speech-like maskers are used, it is expected that the obtained thresholds are worse than predicted by the extended SII model due to additional (*i.e.*, informational) masking.

B. Steepness of the psychometric function

Festen and Plomp (1990) measured entire psychometric functions for speech in stationary and fluctuating noise. Given the larger dynamic range of fluctuating noise, one would expect a larger range in SNR in which the speech is audible, hence a shallower slope for the fluctuating noise masker. Indeed, with normal-hearing subjects, at the level for which a score of 50% is obtained, Festen and Plomp (1990) found a slope of 21.0%/dB and 11.9%/dB for stationary noise and fluctuating noise, respectively. The present Fig. 4, too, shows a shallower slope for fluctuating noise. With the extended SII model, it is possible to predict the slope of the curve obtained with fluctuating noise from that obtained with stationary noise. To that end, it first should be noted that for SNRs from -9 to -1 dB the psychometric curve with stationary noise in Fig. 6 of Festen and Plomp (1990) ranges from 0% to 100%. Figure 4 shows that this SNR range corresponds to a range for the SII of 0.2 to 0.5. An important observation hence is that within the range of 0.2 to 0.5 of the SII, sentence intelligibility changes from 0% to 100%. Within that range for the SII, both curves in Fig. 4 can be well approximated by a linear function. The curve for stationary noise is given by

$$SII_S = (15 + SNR_S)/30, \quad (1)$$

the curve for fluctuating noise is given by

$$SII_F = (27 + SNR_F)/40. \quad (2)$$

Festen and Plomp (1990) describe their curves with a logistic function

$$p(SNR) = \frac{1}{1 + e^{(M - SNR)/S}}, \quad (3)$$

where M is the SNR for which the probability on a correct response $p(SNR)$ is equal to 0.5, and S is the steepness of the function at $p(SNR) = 0.5$. For the stationary noise curve in Fig. 6 of Festen and Plomp (1990), $M = -4.7$ dB and $S = 1.19$ dB (corresponding to 21.0%/dB as given by Festen and Plomp, 1990). For the fluctuating noise curve, $M = -9.7$ dB and $S = 2.10$ dB (corresponding to 11.9%/dB). The data of Fig. 6 of Festen and Plomp (1990) are replotted in Fig. 11, together with the two functions given by Festen and Plomp (1990), given as solid curves. When $SII_S = SII_F$, Eqs. (1) and (2) give the relation between SNR_S and SNR_F

$$SNR_S = (21 + 3SNR_F)/4. \quad (4)$$

By insertion of Eq. (4) into Eq. (3), the shape of the function for fluctuating noise is obtained. This curve is plotted as a dotted line in Fig. 11. The predicted curve for fluctuating noise has a slope of 15.6%/dB and a value for M of -13.3 dB. The curve is about 3.8 dB to the left of the data of Festen and Plomp (1990), but has a slope that fits very well to the data of Festen and Plomp (1990), as can be seen when the curve is shifted 3.8 dB to the right, as has been done in Fig. 11 (dashed curve). The slope fits their data even better than their calculated slope of 11.9%/dB. The fact that the calculated curve does not fall on top of the data of Festen and

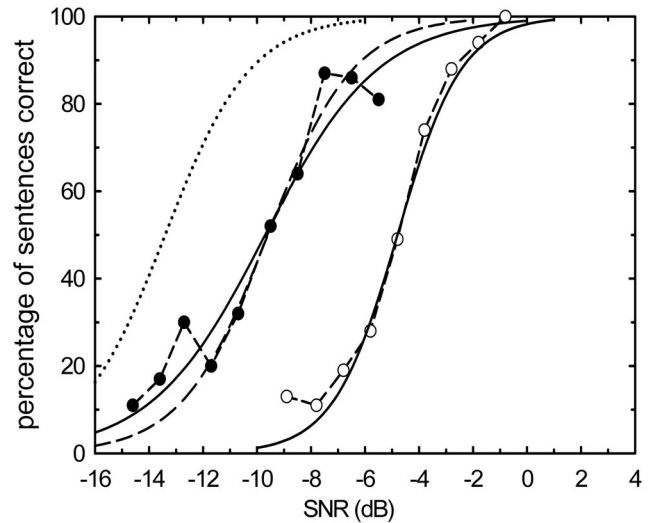


FIG. 11. Percentage of sentences correct as a function of signal-to-noise ratio (dB), for a stationary noise masker (open symbols) and fluctuating noise masker (filled symbols) (replotted from Festen and Plomp, 1990). The two solid curves represent Festen and Plomp's (1990) fit to the data. The dotted curve is predicted by the extended SII model, based on the curve given by Festen and Plomp (1990) for stationary noise. The dashed curve (without symbols) is identical to the dotted curve, except for a shift of 3.8 dB to the right.

Plomp (1990) is due to the fact that Festen and Plomp (1990) shifted their data to the average results.

C. Effect of absolute threshold

With the calculation of the SII, it was assumed that all subjects had normal hearing; that is, thresholds for all frequencies were taken equal to 0 dB(HL). In real life, thresholds deviate to some degree from this value, but with the normal-hearing group it is generally assumed (ANSI S3.6-1996, 1996) that the hearing level is equal to or less than 15 dB(HL). Given the dynamic range of speech (30 dB) and the presentation level of the masking noise, one can calculate the effect of an elevated threshold. With stationary speech noise as a masker, audibility of average conversational speech starts to play a role only at losses of 50 dB(HL) and larger, as can be calculated with the existing SII model. In contrast, with fluctuating noise and interrupted noise, effects become already noticeable at thresholds of 30 or 15 dB(HL), respectively. The effect of hearing loss on the SII is depicted in Fig. 12 for both a stationary noise masker and an interrupted noise masker. As can be seen in this figure, elevating the threshold from 0 to 15 dB(HL) has no effect on the SII with stationary noise, but has a clear effect with interrupted noise. The two curves with interrupted noise start to overlap near an SNR of -15 dB. For the calculations with the extended SII model, little differences in prediction of the SRT in stationary noise were found by variation of the absolute threshold (HL < 50 dB). Figure 12 nevertheless shows that with these fluctuating noise maskers, the effect of absolute threshold can be substantial, especially at lower presentation levels. This could account for the large standard deviation between subjects found by SRT in fluctuating noises (de Laet and Plomp, 1983; Festen, 1987, 1993; Festen and Plomp, 1990; Bronkhorst, 2000; Versfeld and Dreschler, 2002) com-

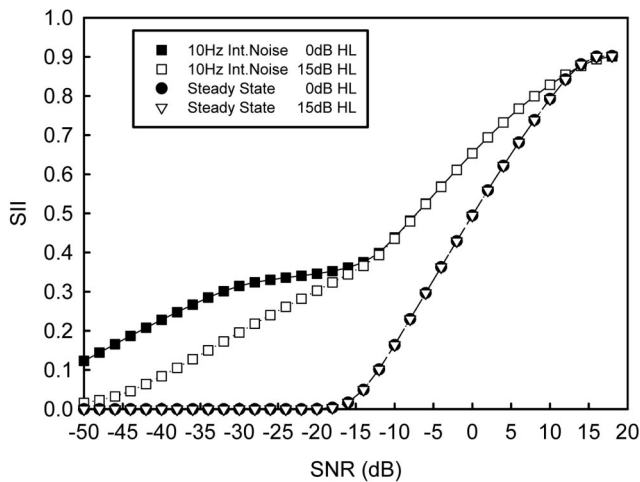


FIG. 12. SII as a function of SNR as calculated with the extended SII model. Filled symbols denote calculations with the absolute threshold set to 0 dB(HL). Open symbols denote calculations with the threshold set to 15 dB(HL). Circles and triangles indicate calculations with a stationary noise masker and squares indicate calculations with interrupted noise masker, respectively, both with the long-term spectrum of the female target speaker. The level of the noises was set to 65 dBA.

pared to the small standard deviation between subjects found by SRT in stationary noises (Plomp and Mimpen, 1979).

D. Effect of window length

With the presentation of the extended SII model, the signals were windowed in time and the length of the time window was frequency dependent. The choice of the time windows was adapted from Moore (1997) and was based on psychophysical data. As discussed above, given these settings, the extended SII model is able to predict the data well. Within a time window, level variations of the signal are averaged. Thus, the longer the time window, the more the signal is smoothed, thus the more the obtained SII will resemble the existing SII (i.e., the SII of stationary noise). On the other hand, if the time windows are taken smaller, all signal variations are caught, which in the case of highly fluctuating maskers as interrupted noise results in better SRTs than actually measured. Calculations have been performed to check whether a single fixed window length for all frequency bands could account for the present data set as well. The results of these calculations show that an optimum fit was obtained with a fixed window length of 12 ms, but that this approach could not account for the data as well as the approach with frequency dependent windows. Yet, it remains possible to manipulate the lengths of the individual time windows, in order to reach an even better fit to the data. However, the present choice of parameters does well, and has the advantage that the window lengths are derived from psychoacoustical measurements. In this paper, rectangular windows have been taken, but future experiments may point to the use of differently shaped windows, such as an exponential window. The latter shape may be more similar to the shape of the forward-masking function.

E. Extensions to the model

In this paper the authors purposely have tried to stay as close as possible to the existing SII model. Extensions to the existing SII model have been proposed, which seem to work well for the SRT with sentences in a given number of noise maskers. To see to what extent the model can be generalized to other types of speech material and noise maskers, measurements should be performed. Although the basic assumptions regarding the extensions may remain valid, it seems plausible that, as with the existing SII model, different speech materials require different weighting functions or window lengths. With the present data set, an SII of 0.35 corresponded to the amount of information required to reach the SRT. These data were obtained with normal-hearing listeners. As discussed extensively by Noordhoek (2000), hearing-impaired subjects often require more speech information to reach threshold, which she attributed to suprathreshold deficits. These deficits probably deal with a decrease in spectral or temporal resolution. With the extended SII model, both decreases in resolution can in principle be modeled by increasing the width of the different frequency bands, or by increasing the window length or window shape. Perhaps more sophisticated adaptations to the SII model [such as the temporal window model of Oxenham (Oxenham and Moore, 1997; Oxenham and Plack, 1997)] are required. It is left to future research to find the extent to which the model is able to describe the data.

F. Other extensions to the SII model

Another shortcoming of the SII model is its inability to account for synergetic and redundant interactions among the various spectral regions of the speech spectrum (Steeneken and Houtgast, 1999; Müsch and Buus, 2001). Due to fact that the SII uses the long-term spectrum of speech and noise (minimum length of 30 s; ANSI S3.5-1997, 1997), these interactions among the various frequency bands are lost. Nevertheless, speech communication is remarkably robust for normal-hearing listeners and does not have to be broadband to be highly intelligible (Allen, 1994; Warren *et al.*, 1995; Lippmann, 1996; Stickney and Assmann, 2000). Steeneken and Houtgast (1999, 2002) implemented a frequency-dependent redundancy correction factor to the STI model, which accounts for synergetic and redundant interactions. Since the STI is related to the SII (van Wijngaarden, 2002), it is in principle possible to implement this redundancy correction factor in the SII calculation method.

V. SUMMARY

The present paper describes an SII-based approach to model SRTs (speech reception thresholds) for sentences masked by fluctuating noise. The basic principle of this approach is that both speech and noise signal are partitioned into small time frames. Within each time frame the instantaneous SII is determined, yielding the speech information available to the listener at that time frame. Next, the SII values of these time frames are averaged, resulting in the SII for that particular noise type. From the literature many SRT values are available for a variety of noise types. In this paper,

it is shown that this approach can give a good account for most existing data. Hence, it forms a valuable extension to the existing SII (ANSI S3.5-1997, 1997) model.

ACKNOWLEDGMENTS

The authors acknowledge Joost Festen and Rob Drullman for providing sound materials of the masking noises used in their papers. Tammo Houtgast, Joost Festen, Gaston Hilkhuisen, and Wouter Dreschler are acknowledged for the stimulating discussions. We are also grateful to Maarten van Beurden, Wouter Dreschler, and Bas Franck for their comments on earlier versions of this paper. We thank the associate editor, Peter Assmann, and the two anonymous reviewers for their detailed constructive comments. Finally, Jan Koopman and especially László Körössy are acknowledged for their help with the computer programming.

- Allen, J. B. (1994). "How do humans process and recognize speech," *IEEE Trans. Speech Audio Process.* **2**, 567–577.
- ANSI (1996). ANSI S3.6-1996, "American National Standard Methods for Specification for audiometers" (American National Standards Institute, New York).
- ANSI (1997). ANSI S3.5-1997, "American National Standard Methods for Calculation of the Speech Intelligibility Index" (American National Standards Institute, New York).
- Bacon, S. P., Opie, J. M., and Montoya, D. Y. (1998). "The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds," *J. Speech Lang. Hear. Res.* **41**, 549–563.
- Bosman, A. J., and Smoorenburg, G. F. (1995). "Intelligibility of Dutch CVC syllables and sentences for listeners with normal hearing and with three types of hearing impairment," *Audiology* **34**, 260–284.
- Bronkhorst, A. W. (2000). "The Cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," *Acustica* **86**, 117–128.
- Bronkhorst, A. W., and Plomp, R. (1992). "Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing," *J. Acoust. Soc. Am.* **92**, 3132–3139.
- Brungart, D. S. (2001). "Informational and energetic masking effects in the perception of two simultaneous talkers," *J. Acoust. Soc. Am.* **109**, 1101–1109.
- Brungart, D. S., and Simpson, B. D. (2002). "The effects of spatial separation in distance on the informational and energetic masking of a nearby speech signal," *J. Acoust. Soc. Am.* **112**, 664–676.
- Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. (2001). "Informational and energetic masking effects in the perception of multiple simultaneous talkers," *J. Acoust. Soc. Am.* **110**, 2527–2538.
- Carhart, R., Tillman, T. W., and Greetis, E. S. (1969). "Perceptual masking in multiple sound backgrounds," *J. Acoust. Soc. Am.* **45**, 694–703.
- de Laat, J. A. P. M., and Plomp, R. (1983). "The reception threshold of interrupted speech for hearing-impaired listeners," in *Hearing—Physiological Bases and Psychophysics*, edited by R. Klinke and R. Hartman (Springer, Berlin), pp. 359–363.
- Dirks, D. D., Bell, T. S., Rossman, R. N., and Kincaid, G. E. (1986). "Articulation index predictions of contextually dependent words," *J. Acoust. Soc. Am.* **80**, 82–92.
- Drullman, R., and Bronkhorst, A. W. (2000). "Multichannel speech intelligibility and talker recognition using monaural, binaural, and three-dimensional auditory presentation," *J. Acoust. Soc. Am.* **107**, 2224–2235.
- Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2002). "Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing," *J. Acoust. Soc. Am.* **111**, 2897–2907.
- Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2003). "Recovery from prior stimulation: Masking of speech by interrupted noise for younger and older adults with normal hearing," *J. Acoust. Soc. Am.* **113**, 2084–2094.
- Duquesnoy, A. J. (1983). "Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons," *J. Acoust. Soc. Am.* **74**, 739–743.
- Eddins, D. A., Hall, III, J. W., and Grose, J. H. (1992). "The detection of temporal gaps as a function of frequency region and absolute noise bandwidth," *J. Acoust. Soc. Am.* **91**, 1069–1077.
- Festen, J. M. (1987). "Speech-perception threshold in a fluctuating background sound and its possible relation to temporal resolution," in *The Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Martinus Nijhoff, Dordrecht), pp. 461–466.
- Festen, J. M. (1993). "Contributions of comodulation masking release and temporal resolution to the speech-reception threshold masked by an interfering voice," *J. Acoust. Soc. Am.* **94**, 1295–1300.
- Festen, J. M., and Plomp, R. (1990). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing," *J. Acoust. Soc. Am.* **88**, 1725–1736.
- Fletcher, H., and Galt, R. H. (1950). "The perception of speech and its relation to telephony," *J. Acoust. Soc. Am.* **22**, 89–151.
- French, N. R., and Steinberg, J. C. (1947). "Factors governing the intelligibility of speech sounds," *J. Acoust. Soc. Am.* **19**, 90–919.
- Glasberg, B. R., and Moore, B. C. (1992). "Effects of envelope fluctuations on gap detection," *Hear. Res.* **64**, 81–92.
- Gustafsson, H. A., and Arlinger, S. D. (1994). "Masking of speech by amplitude-modulated noise," *J. Acoust. Soc. Am.* **95**, 518–529.
- Hogan, C. A., and Turner, C. W. (1998). "High-frequency audibility: Benefits for hearing-impaired listeners," *J. Acoust. Soc. Am.* **104**, 432–441.
- Houtgast, T., Steeneken, H. J., and Bronkhorst, A. W. (1992). "Speech communication in noise with strong variations in the spectral or the temporal domain," *Proceedings of the 14th International Congress on Acoustics*, Vol. 3, pp. H2–6.
- Howard-Jones, P. A., and Rosen, S. (1992). "The perception of speech in fluctuating noise," *Acustica* **78**, 258–272.
- Howard-Jones, P. A., and Rosen, S. (1993). "Unmodulated glimpsing in checkerboard noise," *J. Acoust. Soc. Am.* **93**, 2915–2922.
- Hygge, S., Ronnberg, J., Larsby, B., and Arlinger, S. (1992). "Normal-hearing and hearing-impaired subjects' ability to just follow conversation in competing speech, reversed speech, and noise backgrounds," *J. Speech Hear. Res.* **35**, 208–215.
- Kamm, C. A., Dirks, D. D., and Bell, T. S. (1985). "Speech recognition and the Articulation Index for normal and hearing-impaired listeners," *J. Acoust. Soc. Am.* **77**, 281–288.
- Kryter, K. D. (1962a). "Methods for the calculation and use of the articulation index," *J. Acoust. Soc. Am.* **34**, 1689–1697.
- Kryter, K. D. (1962b). "Validation of the articulation index," *J. Acoust. Soc. Am.* **34**, 1698–1702.
- Larsby, B., and Arlinger, S. (1994). "Speech recognition and just-follow-conversation tasks for normal-hearing and hearing-impaired listeners with different maskers," *Audiology* **33**, 165–176.
- Licklider, J. C. R., and Guttman, N. (1957). "Masking of speech by line-spectrum interference," *J. Acoust. Soc. Am.* **29**, 287–296.
- Lippmann, R. P. (1996). "Accurate consonant perception without mid-frequency speech energy," *IEEE Trans. Speech Audio Process.* **4**, 567–577.
- Middelweerd, M. J., Festen, J. M., and Plomp, R. (1990). "Difficulties with speech intelligibility in noise in spite of a normal pure-tone audiogram," *Audiology* **29**, 1–7.
- Miller, G. A. (1947). "The masking of speech," *Psychol. Bull.* **44**, 105–129.
- Miller, G. A., and Licklider, J. C. R. (1950). "The intelligibility of interrupted speech," *J. Acoust. Soc. Am.* **22**, 167–173.
- Moore, B. C. (1997). *An Introduction to the Psychology of Hearing*, 4th ed. (Academic, London).
- Moore, B. C., Peters, R. W., and Glasberg, B. R. (1996). "Detection of decrements and increments in sinusoids at high overall levels," *J. Acoust. Soc. Am.* **99**, 3669–3677.
- Müsch, H., and Buus, S. (2001). "Using statistical decision theory to predict speech intelligibility. II. Measurement and prediction of consonant-discrimination performance," *J. Acoust. Soc. Am.* **109**, 2910–2920.
- Neijenhuis, K., Sink, A., Priester, G., van Kordenoordt, S., and van der Broek, P. (2002). "Age effects and normative data on a Dutch test battery for auditory processing disorders," *Int. J. Audiol.* **41**, 334–346.
- Nelson, P. B., Jin, S. H., Carney, A. E., and Nelson, D. A. (2003). "Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners," *J. Acoust. Soc. Am.* **113**, 961–968.
- Nilsson, M., Soli, S. D., and Sullivan, J. A. (1994). "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise," *J. Acoust. Soc. Am.* **95**, 1085–1099.
- Noordhoek, I. M. (2000). "Intelligibility of narrow-band speech and its

- relation to auditory functions in hearing-impaired listeners," Doctoral thesis, Free University, Amsterdam.
- Oxenham, A. J., and Moore, B. C. (1994). "Modeling the additivity of nonsimultaneous masking," *Hear. Res.* **80**, 105–118.
- Oxenham, A. J., and Moore, B. C. (1997). "Modeling the effects of peripheral nonlinearity in normal and impaired hearing," in *Modeling Sensorineural Hearing Loss*, edited by W. Jesteadt (Erlbaum, Mahwah, NJ), pp. 273–288.
- Oxenham, A. J., and Plack, C. J. (1997). "A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing," *J. Acoust. Soc. Am.* **101**, 3666–3675.
- Oxenham, A. J., Rosengard, P. S., and Braid, L. D. (2004). "Perceptual consequences of normal and abnormal peripheral compression: Potential links between psychoacoustics and speech perception," *J. Acoust. Soc. Am.* **115**, 2421.
- Pavlovic, C. V. (1984). "Use of the articulation index for assessing residual auditory function in listeners with sensorineural hearing impairment," *J. Acoust. Soc. Am.* **75**, 1253–1258.
- Pavlovic, C. V. (1987). "Derivation of primary parameters and procedures for use in speech intelligibility predictions," *J. Acoust. Soc. Am.* **82**, 413–422.
- Pavlovic, C. V., and Studebaker, G. A. (1984). "An evaluation of some assumptions underlying the articulation index," *J. Acoust. Soc. Am.* **75**, 1606–1612.
- Pavlovic, C. V., Studebaker, G. A., and Sherbecoe, R. L. (1986). "An articulation index based procedure for predicting the speech recognition performance of hearing-impaired individuals," *J. Acoust. Soc. Am.* **80**, 50–57.
- Peters, R. W., Moore, B. C., and Baer, T. (1998). "Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people," *J. Acoust. Soc. Am.* **103**, 577–587.
- Plack, C. J., and Oxenham, A. J. (1998). "Basilar-membrane nonlinearity and the growth of forward masking," *J. Acoust. Soc. Am.* **103**, 1598–1608.
- Plomp, R. (1964). "Rate of decay of auditory sensation," *J. Acoust. Soc. Am.* **36**, 277–282.
- Plomp, R., and Mimpen, A. M. (1979). "Improving the reliability of testing the speech reception threshold for sentences," *Audiology* **18**, 43–52.
- Rankovic, C. M. (1998). "Factors governing speech reception benefits of adaptive linear filtering for listeners with sensorineural hearing loss," *J. Acoust. Soc. Am.* **103**, 1043–1057.
- Rankovic, C. M. (2002). "Articulation index predictions for hearing-impaired listeners with and without cochlear dead regions," *J. Acoust. Soc. Am.* **111**, 2545–2548.
- Shailer, M. J., and Moore, B. C. (1983). "Gap detection as a function of frequency, bandwidth, and level," *J. Acoust. Soc. Am.* **74**, 467–473.
- Shailer, M. J., and Moore, B. C. (1987). "Gap detection and the auditory filter: Phase effects using sinusoidal stimuli," *J. Acoust. Soc. Am.* **81**, 1110–1117.
- Smoorenburg, G. F. (1992). "Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram," *J. Acoust. Soc. Am.* **91**, 421–437.
- Steeneken, H. J. (1992). "On measuring and predicting speech intelligibility," Doctoral thesis, University of Amsterdam.
- Steeneken, H. J., and Houtgast, T. (1980). "A physical method for measuring speech-transmission quality," *J. Acoust. Soc. Am.* **67**, 318–326.
- Steeneken, H. J., and Houtgast, T. (1999). "Mutual dependence of the octave-band weights in predicting speech intelligibility," *Speech Commun.* **28**, 109–123.
- Steeneken, H. J., and Houtgast, T. (2002). "Validation of the revised STI method," *Speech Commun.* **38**, 413–425.
- Stickney, G. S., and Assmann, P. F. (2001). "Acoustic and linguistic factors in the perception of bandpass-filtered speech," *J. Acoust. Soc. Am.* **109**, 1157–1165.
- Studebaker, G. A., Pavlovic, C. V., and Sherbecoe, R. L. (1987). "A frequency importance function for continuous discourse," *J. Acoust. Soc. Am.* **81**, 1130–1138.
- Studebaker, G. A., Taylor, R., and Sherbecoe, R. L. (1994). "The effect of noise spectrum on speech recognition performance-intensity functions," *J. Speech Hear. Res.* **37**, 439–448.
- ter Keurs, M., Festen, J. M., and Plomp, R. (1993). "Limited resolution of spectral contrast and hearing loss for speech in noise," *J. Acoust. Soc. Am.* **94**, 1307–1314.
- Trine, T. D. (1995). "Speech recognition in modulated noise and temporal resolution: Effects of listening bandwidth," Unpublished doctoral dissertation, University of Minnesota, Twin Cities, MN.
- Turner, C. W., and Henry, B. A. (2002). "Benefits of amplification for speech recognition in background noise," *J. Acoust. Soc. Am.* **112**, 1675–1680.
- van Wijngaarden, S. J. (2002). "Past, Present and future of the speech transmission index," Proceedings of the International Symposium on STI, TNO Human Factors (Soesterberg, The Netherlands).
- van Wijngaarden, S. J. (2003). "The intelligibility of non-native speech," Doctoral thesis, Free University, Amsterdam.
- Versfeld, N. J., and Dreschler, W. A. (2002). "The relationship between the intelligibility of time-compressed speech and speech in noise in young and elderly listeners," *J. Acoust. Soc. Am.* **111**, 401–408.
- Versfeld, N. J., Daalder, L., Festen, J. M., and Houtgast, T. (2000). "Method for the selection of sentence materials for efficient measurement of the speech reception threshold," *J. Acoust. Soc. Am.* **107**, 1671–1684.
- Warren, R. M., Riener, K. R., Bashford, J. A., and Brubaker, B. S. (1995). "Spectral redundancy: Intelligibility of sentences heard through narrow spectral slits," *Percept. Psychophys.* **57**, 175–182.

Perception of pitch location within a speaker's F0 range^{a)}

Douglas N. Honorof^{b)} and D. H. Whalen

Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511

(Received 24 March 2004; revision received 2 November 2004; accepted 5 November 2004)

Fundamental frequency (F0) is used for many purposes in speech, but its linguistic significance is based on its relation to the speaker's range, not its absolute value. While it may be that listeners can gauge a specific pitch relative to a speaker's range by recognizing it from experience, whether they can do the same for an unfamiliar voice is an open question. The present experiment explored that question. Twenty native speakers of English (10 male, 10 female) produced the vowel /a/ with a spoken (not sung) voice quality at varying pitches within their own ranges. Listeners then judged, without familiarization or context, where each isolated F0 lay within each speaker's range. Correlations were high both for the entire range (0.721) and for the range minus the extremes (0.609). Correlations were somewhat higher when the F0s were related to the range of all the speakers, either separated by sex (0.830) or pooled (0.848), but several factors discussed here may help account for this pattern. Regardless, the present data provide strong support for the hypothesis that listeners are able to locate an F0 reliably within a range without external context or prior exposure to a speaker's voice. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1841751]

PACS numbers: 43.71.An, 43.71.Bp, 43.71.Es [PFA].

Pages: 2193–2200

I. INTRODUCTION

Fundamental frequency (F0) carries information about many different linguistic and paralinguistic aspects of the speech signal, but it does so in a speaker-dependent way. A pitch-accent system like that of Japanese, for example, distinguishes words on the basis of which mora bears a shift in relative F0. However, F0 height is not specified by the language in general; rather, what counts as high or low varies by speaker (Leather, 1983; Moore and Jongman, 1997). Irrespective of which parameters require F0 manipulation in a given phonology, the most obvious source of difference in pitch is speaker sex. Peterson and Barney (1952) found a difference of approximately three-quarters of an octave between adult males and adult females in average F0s of citation-form monosyllables across vowels, and approximately a full octave's difference when comparing analogous F0 measures for adult males and children. It appears, then, that "high pitch" cannot be specified as a particular F0 value, because any value (if attainable) could be high for some speakers and low for others.

How, then, do listeners gauge where within the speaker's range a particular pitch lies? One possibility, of course, is that direct exposure to a voice over time allows a listener to learn the range of pitches that a speaker can and does produce. If so, the perception of linguistic pitch targets may become possible as a function of exposure to a speaker's voice. Indeed, there is evidence that familiarity with a voice is a factor in speech perception. For example, words spoken by familiar voices are easier to recognize (Palmeri *et al.*, 1993) and identify in noise (Nygard and Pisoni, 1998). Such effects of familiarity in perception are not surprising given that linguistic and sociophonetic features are intertwined

with speaker-indexical information in the speech signal (Ladefoged and Broadbent, 1957; Fellowes *et al.*, 1997; Remez *et al.*, 1997) and in word memory (Goldinger, 1996). It may be that listeners perform "speaker normalization" in order to strip away linguistically extraneous information from the linguistic message, but it is at least as likely, as recent findings suggest, that listeners use a normalization strategy in order to disentangle invariant phonological information and invariant speaker information from each other without necessarily forgetting the latter (see Sheffert and Fowler, 1995). There may remain much that we do not yet understand about how speaker normalization is accomplished by human listeners (e.g., Johnson and Mullennix, 1997), but we should be careful not to assume that perception always entails a computational process of normalization. Clearly, listeners successfully perceive the speech of unfamiliar voices—voices not stored in memory. In such cases, no exposure-based normalization strategy is available to the listener.

It is unclear whether previous exposure to a voice is required for accurate pitch perception. When a sentential context is given, listeners can estimate the location of a pitch within a speaker's range (Wong and Diehl, 2003). Other studies suggest success with contexts even shorter than a sentence. For example, in some languages, there are utterances as short as a single syllable that span a large region of the speaking range as in the case of the falling tone of standard Mandarin, in which case exposure to a single utterance of a single tone might, conceivably, provide enough information about a speaker's range to allow listeners to normalize F0. Indeed, a study by Moore and Jongman (1997) indicates that perception of a Mandarin tone having an F0 turning point or change in F0 does not depend on context, thus dynamic F0 is presumably calibrated syllable internally. If prior exposure to a short contour is required for the listener to succeed at normalizing pitch, logically, the listener should

^{a)}Portions of this work were presented at the 146th Meeting of the Acoustical Society of America, Austin, Texas, November 2003.

^{b)}Electronic mail: honorof@haskins.yale.edu

not be able to normalize pitch correctly if presented only with *steady* F0s. It remains to be seen whether listeners can correctly assess the location of F0 within the spoken pitch range (*tessiture*) of an unfamiliar individual on the basis of isolated *natural* utterances without first having been exposed to F0 contours within the speaker's range.

If perception of steady F0 does not require prior exposure to the speaker's range, we must determine how the listener is able to accomplish the task. One basis for pitch location within an F0 range might be voice quality. The literature has produced a number of differing claims regarding the nature of dependencies between voice quality (as reflected in glottal pulse characteristics) and F0. Common measures of glottal pulse characteristics include overall intensity, the open quotient of the voice source (that is, the relative time during which the glottis remains open during a pitch period), skewness of the glottal pulse, the strength of the relationship between the amplitudes of F0 and $F0^2$ (that is, $H2-H1$), etc. Swerts and Veldhuis (2001) review this literature and present the results of a highly constrained study of their own aimed at exploring the proposal that listeners "might be able to locate an utterance in a speaker's speech range on the basis of voice quality (298)." They conclude that there is some evidence that the perception of F0 depends on voice quality, but do not go so far as to explore directly the ramifications of such a finding for perception (see also Fant *et al.*, 1985; Klatt and Klatt, 1990; Gobl and Ní Chasaide, 1992). Di Paolo and Faber (1990) have reported that voice quality (breathy versus creaky phonation as indexed by $F1-F0$) can be used by listeners to distinguish tense from lax vowels. Taken as a whole, these studies make direct assessment of voice quality a possible means for the immediate normalization of an individual speaker's pitch without prior exposure to a broad range of productions by the voice of the individual. Therefore, we hypothesize that listeners should be able to use immediately available information about voice quality to locate where within a speaker's range the pitch of a vowel lies, even when presented with unfamiliar voices and virtually no context for their judgments.

If listeners are unable to perform this task, we will have to conclude that the perception of pitch does indeed require contextualization or familiarization with a particular speaker's pitch range. If, on the other hand, as predicted, listeners are able to successfully identify the relative location of isolated steady pitches without prior exposure to a speaker's voice (and thus to a speaker's *tessiture*), we will have evidence that acoustic parameters of some sort must co-vary with F0 so as to make an immediate identification of relative F0 possible.

II. METHODS

A. Materials

1. Speakers

Twenty native speakers of North American English (10 men, 10 women) participated in the production phase of the experiment. Ages ranged from 20 to 78. Each speaker passed a prescreening administered by questionnaire. The answers of those who participated in the production phase of the ex-

periment indicated that they had neither training nor experience as a vocal soloist, that they were nonsmokers (for at least the preceding five years), that they were in good vocal health (including the absence of symptoms of congestion, coughing, post-nasal drip, or fatigue on the day of the test), that they were not prone to severe or chronic respiratory allergies, and that they had never received the diagnosis of a communication disorder, vocal-fold pathology, or neuromotor impairment. Furthermore, on the day of the test, all 20 speakers passed manual binaural audiometric screening via a Beltone 110 in a sound-attenuated chamber. The audiometric tests demonstrated hearing-threshold levels at audiometric (pulsed) frequencies between 0.125 and 1 kHz at 25 dB or lower—well below the average amplitude of conversational speech.

2. Stimulus generation

In order to determine which F0s we should attempt to elicit from a given speaker and in order to allow us, at a later date, to calculate where those F0s lay within speaker-specific voice ranges, we elicited spoken *tessitures* (vocal ranges) from each speaker. In a double-isolated sound booth, the speaker was instructed to speak, not sing (Titze, 1995), /a/ glissandos as though they were saying the words at a conversational "loudness" to the experimenter, who was seated in the booth approximately one meter away from the speaker. The speaker spoke an /a/ by beginning at a comfortable (self-determined) habitual pitch (see Fairbanks, 1960), then increasing fundamental frequency continuously until modal phonation could no longer be sustained. Multiple rising glissandos of this type were elicited from each speaker until the experimenter was satisfied that he had recorded one smoothly rising, nonsung glissando that included the extreme high end of the speaker's range. Next, a set of falling /a/ glissandos was elicited in a similar manner beginning at a speaker-specific habitual pitch and falling until the speaker was no longer able to sustain phonation at the low end of the range. The number of attempts was typically three but varied from two to eight. During elicitation of spoken glissandos, the speaker was asked to begin again if he or she sensed, or the experimenter judged, that the voice had taken on a sung quality or if the audio signal was clipped or if a vowel token was produced noticeably more quietly than other tokens.

Speaker-specific ranges were calculated from these glissandos using autocorrelation in Praat v4.1 (range=45–900 Hz; 15 maximum candidates), in particular, to provide information about the high end of the *tessiture*. Autocorrelation was supplemented by interactive measurement of the reciprocal of the arithmetic mean of three successive pitch periods. The floor of the spoken *tessiture* was defined as the basal pitch, specifically, the frequency at which periodic F0 was no longer analyzable by autocorrelation or, in dubious cases, identified by direct measurement of three pitch periods. The ceiling of the spoken *tessiture* was defined as the frequency of the first falsetto break extracted from the rising glissando during which nonfalsetto voice quality was sustained at the highest frequency. In clinical practice, falsetto is often included in a normative voice range profile (VRP, also called the " F_0-SPL_{min} voice profile," " F_0/SPL profile,"

“voice field,” “voice area,” or “phonetogram”) in an attempt to assess voice disorders (especially of singers), in which case the falsetto break is often seen as a small notch in intensity (<5 dB) on the loud side of the (normally) oval shape of the profile (Schutte and Seidner, 1983; Heylen *et al.*, 2002). Although falsetto may be used in speech, not just in singing, we excluded falsetto from our speaker-specific ranges in order to increase the difficulty of the ranking task for the listeners (see later). Because we were not interested in the margins of the F0 range, no effort was made to control for variations in extrema within speakers as a function of time of day (see Gelfer, 1989), beyond the collection of multiple ranges from each speaker.

Once speaker-specific ranges had been established, each speaker was asked to sustain a spoken /a/ on at least eight distinct notes spaced roughly evenly within his or her nonfalsetto spoken F0 range. The instruction to speak rather than sing /a/ was met with surprisingly little resistance, perhaps because nonsingers had been recruited by design, some of whom having expressed relief that they were not expected to sing. As was the case with the glissandos, occasional lapses into a sung voice quality also occurred in the production of steady tones, and sung tokens were replaced. In order to help these nonsingers space notes out within their own ranges, they were provided with sinusoidal audio stimuli. Specifically, they heard four 300 ms pure tones per octave (notes c, e, g, and a from 49 to 524 Hz) in which a linear ramp function was applied to the amplitude over the first 100 ms and over the last 100 ms. The experimenters selected at least eight such auditory stimuli within the speaker’s tessiture and presented them over a harman/kardon multimedia speaker system with adjustable output levels. Output was always well above measured auditory threshold levels for the speaker. The talker was asked to respond to each stimulus by attempting to match the pitch of the tone while speaking /a/ at a conversational “loudness.” In order to help the speaker further, we used a Kay Visipitch II running in DOS to provide real-time visual feedback on success of pitch-matching. The participant saw a trace of his or her own pitch along with horizontal guide lines placed at 10 Hz above and 10 Hz below the target pitch. Speakers were asked to hold each tone for four to five seconds, then to inspire (to encourage resetting of the glottis) and to produce at least another two vowels, also separated by a breath. Vowels so produced were also digitized as 44.1 kHz, 16-bit NSP sound files using a near-measurement-quality microphone (Earthworks QTC1) routed to the Visipitch via a hardware high-pass filter (Focusrite ISA220, cutoff=30 Hz, 18 dB/octave rolloff) under speaker-independent input gain. We excluded tokens with noticeable F0 or intensity excursions, as seen in acoustic analysis displays. For each speaker, the eight tokens that were best spread out throughout the speaker’s range were then selected. For each item selected, the first 500 ms containing no noticeable F0 or intensity excursions were extracted from the longer vowel. A linear amplitude ramp was applied to the first and last 100 ms of this extracted segment.

B. Listeners

Twelve native speakers of North American English (6 men, 6 women) participated in the listening task. Ages ranged from 20 to 58. Each listener was asked to respond to a prescreening questionnaire. Answers indicated that none had ever been diagnosed with a communication disorder. Furthermore, on the day of the test, all 12 listeners passed a manual binaural audiometric screening via a Beltone 110 in a sound-attenuated chamber. The audiometric tests demonstrated hearing threshold levels at audiometric (pulsed) frequencies between 0.125 and 1 kHz at 30 dB or lower, well below the lowest amplitude present in the stimulus materials.

C. Procedures

For each stimulus we had recorded, we calculated the achieved F0—not the target F0—as a percentage of the corresponding speaker-specific nonfalsetto tessiture using the formula in Eq. (1), where $F0_{\text{ratio}}$ is the percentage of the individual speaker’s tessiture with which we aimed to correlate listener rankings, $F0_{\text{ach}}$ is the fundamental frequency location actually achieved by the speaker, T_{base} is the lowest sustainable F0 within the speaker’s spoken tessiture, and T_{range} is the tessiture itself.

$$F0_{\text{ratio}} = \frac{F0_{\text{ach}} - T_{\text{base}}}{T_{\text{range}} \cdot 0.01}. \quad (1)$$

Tessitures were based on spoken glissandos, not steady F0s, but highest speakable steady F0s and the top end of the nonfalsetto spoken glissando ranges differed slightly, as reported elsewhere (Reich *et al.*, 1990, but see Zraick *et al.*, 2000). Therefore, for two female speakers, the highest F0 values lay above the speaker’s tessiture, 3% and 14% above for one speaker, and 5% above for the other. Half of the lowest male tokens and half of the lowest female tokens, though they fell within the tessiture, were produced with highly periodic glottal fry. Two instances of the second lowest tones (one male, one female) were also produced with fry. Tessitures for each speaker appear coded by sex in Fig. 1.

Listeners were warned that they would be hearing the voices of different speakers using different parts of their speaking ranges, and asked to identify where within the speaker’s range each token fell. The list of stimuli, consisting of one token of each of eight F0s collected from each speaker ($8 \times 20 = 160$ stimuli), was randomized across speakers and tokens separately for each listener and presented diotically one at a time over headphones. Stimuli were presented at a comfortable output level set by the experimenter and kept constant across listeners. Responses were solicited via a custom MATLAB graphical user interface (GUI) on a flat-panel display. The GUI window accepted input via clickable buttons numbered sequentially and arranged vertically, with the word “high” at the top and “low” at the bottom. Listeners used these buttons to indicate the location of an /a/ token within a speaker’s pitch range on a nine-point scale numbered “1” (“just about as low as they could go”) to “9” (“just about as high as they could go”). A nine-point scale was used so that the eight F0 levels would not be underrepresented. Listeners were instructed to choose “5” if they

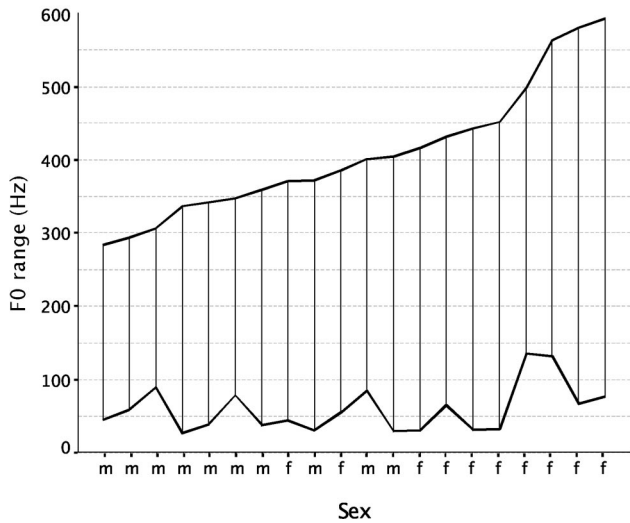


FIG. 1. Line chart of tessitures by speaker with high and low values connected. The chart is arranged from lowest falsetto break (left) to highest (right). Male falsetto breaks cluster toward the low end and gradually rise toward the female falsetto breaks with little overlap. Sex-based differences in the location of the low end of the tessiture are less striking. There, ranges bottom out in the high 20s and low 30s for the lowest voices in both sexes, though the 3 highest points are female (594, 563, 581, and 563 Hz) and are measured at over 150 Hz higher than the highest male low end.

thought the pitch was “just normal,” but listeners were told that they would probably need to use all the numbers equally often. They were able to press a “repeat” button to hear a stimulus again, but they were discouraged from listening more than twice. They were also allowed to rest at any point if they wished.

After the task had been completed, listeners were debriefed regarding the perceived difficulty of the task and strategies they believed they had employed. Additionally, all were asked to confirm that they had correctly understood the instruction to indicate where each token lay within its speaker-specific range. All of them did so.

III. RESULTS

A. Correlations by speaker tessitures

A linear regression was run on two variables, with assigned *ranking* of individual tokens (12 listeners) as the dependent variable and, as the independent variable, *F0 location* of the same tokens (calculated as a percentage of the relevant speaker-specific tessiture based on a log base 10 transformation of the high and low F0s from the glissandos). In the scatterplot given as Fig. 2, a relatively dense cluster (a binned range of values) of individual tokens appears as a collection of densely petalled “sunflowers” surrounding a regression line. Correlations were calculated for each listener separately. In all cases r was positive, ranging from 0.587 to 0.789, and $p < 0.001$. In order to render the variance more independent of the population, Fisher’s z transformation was applied to each r (Fisher, 1915). The mean of the resulting z values was calculated and that mean inverted into an r ; this value was 0.721 and indicates a relationship between assigned rankings and location of the tones within speaker-specific tessitures.

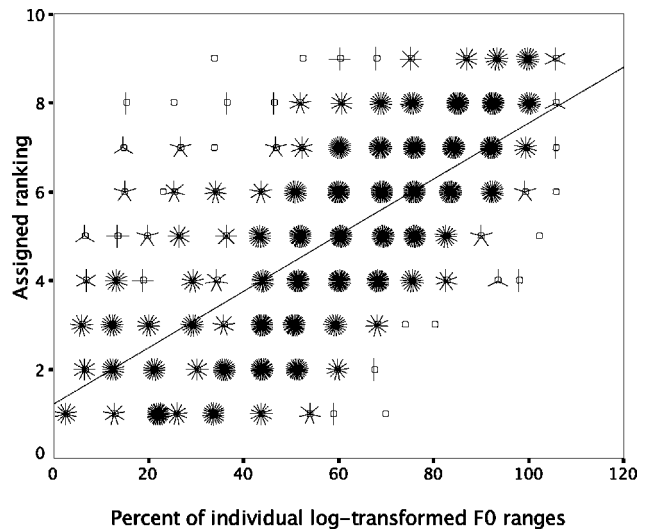


FIG. 2. Sunflower scatterplot with regression line between rankings assigned by 12 listeners and log base 10 transformed locations of eight spoken F0s per speaker (20 speakers). Locations are reported as a percentage of speaker-specific spoken F0 ranges. Vowel: /a/. A direct linear relationship is apparent. Each sunflower represents one or more cases that occur very close together on the scatterplot, with each case represented as a sunflower petal. Each sunflower is positioned at the mean of the cluster of cases that it represents. Each dimension of the sunflower cell is 1/15 of the appropriate range.

B. Correlations by speaker tessitures, extrema excluded

In order to exclude the possibility that the correlation reported above was skewed by listener reliance on unusual voice qualities located at range extrema, we ran a further linear regression between assigned rankings from the middle six F0s for each speaker and actual locations of the middle six F0s within speaker-specific tessitures. That is, the high and low F0s for each talker were excluded. Half of these low values were judged by the experimenters to have a creaky voice quality. (Two other tokens that were second lowest were judged to have a creaky quality.) Here also, the scatterplot (see Fig. 3) suggests a linear relationship between ranking and location, and all r values were positive and significant ($p < 0.001$); they ranged from 0.463 to 0.716. The inverse Fisher mean z (calculated as above) was 0.609, which indicates a direct relationship between the variables, though it suggests a weaker relationship than reported above for all eight F0s per speaker. However, a reduction in the strength of the correlation does not necessarily indicate that information around the center of each speaker’s range was not sufficient for the listener. We return to this point in the Discussion.

C. Correlations by generic ranges

To test for the possibility that listeners were actually locating F0 relative to F0 ranges heard elsewhere (and perhaps separately for the two sexes), two further analyses were performed. One assumed that listeners would be able to determine for each stimulus whether a male or a female had uttered the vowel. Listeners are good at doing this based on sentence-length stretches of speech (Lass *et al.*, 1978), and on short segments of speech (Ingemann, 1968). It is also

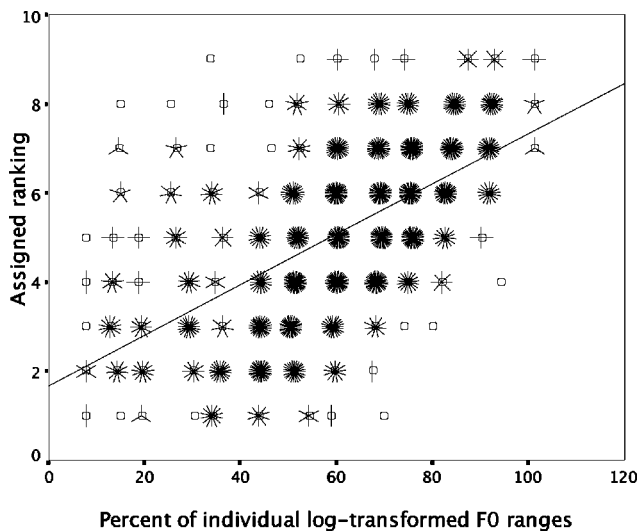


FIG. 3. Sunflower scatterplot with regression line between rankings assigned by 12 listeners and log base 10 transformed locations of six spoken tones per speaker (20 speakers) with locations reported as a percentage of speaker-specific spoken F0 ranges. The single highest and lowest tone for each speaker have been excluded from the regression, along with corresponding rankings. Vowel: /a/. A direct linear relationship is apparent. Plotting dimensions are as in Fig. 2.

possible to reliably distinguish sex on the basis of the acoustics of vowel segments alone (Childers and Wu, 1991). All of these studies use central F0 values, so it is not clear whether the current listeners would be able to make this distinction throughout the range. Nonetheless, it seemed reasonable to assume that such a distinction might be of perceptual value, so a correlation was calculated as before but using the male or female extreme values as the definition of the range rather than the individual speaker's extremes.

The pattern of correlations is very similar to that found before, so no figure is given. The correlations for the sex-linked estimates were all positive and significant ($p < 0.001$), ranging from 0.611 to 0.901, and the inverse Fisher mean z was again positive (0.830, $p < 0.001$). The magnitude of the correlation was somewhat larger than that for the individual tessitures.

A final analysis used the overall extreme values (low from the lowest male, high from the highest female) as defining the global "tessiture." The correlation coefficients for the global, non-sex-linked range were also positive and significant ($p < 0.001$), ranging from 0.611 to 0.934, with an inverse Fisher mean z of 0.848. These correlations are similar to those found in van Bezoooyen (1984) for mean F0 and the perceived pitch level of her natural speech stimuli (expressing nine emotions plus a neutral reading).

D. Comparison of the two types of ranges

In order to assess the fit between the estimate of the location within range and the ranges based on individual and group tessitures, we have plotted the eight sustained F0s collected from each speaker, superimposing speaker-specific upper and lower tessiture boundaries elicited from spoken glissandos, plotting data according to the value of the bottom of the tessiture (Fig. 4). Colors in the figure represent a continuous gradation based on averaged listener rankings. The figure

allows some comparison of success with the two types of range. Speaker-specific patterns of color updrift across the figure are balanced by global patterns seen in the relative stability of the mid-range bands of color across speakers.

Colored dots tend to expand to fill the tessitures whose boundaries are demarcated by black lines, though there are notable gaps between the lowest or highest F0 attained and the limit of the speaker-specific tessiture. That is, there are steady F0s that were produced and excluded either because eight values had already been obtained from that speaker or because no acceptable token was obtained by our elicitation technique. Those gaps are seen especially, on the low end, in the productions of speakers 6 and 8 among the males and 4, 11, 12, and 17 among the females, and on the high end, in the productions of speakers 2, 3, 6, 9, 18, and 20 among the males and 15 and 16 among the females. The lack of a dark blue dot for the lowest production of speakers 7, 9, 13, and 15 is more consistent with a global interpretation of F0 range fitting, since these attained F0s were close enough to the lower limit of the tessitures that listeners ought to have been able to assign them a ranking of 1 if they were able to locate F0 within individual ranges. Here, just as the correlation analyses show, there is evidence for both specific and general patterns.

IV. DISCUSSION

Significant positive correlations obtain between the assigned ranking and the actual location of tones within individually defined tessitures. However, similar correlations were also found when the F0s were scaled to the ranges of the speakers as a group. Since there was sizable overlap among the F0 ranges of our current talkers, it was not possible to distinguish which factor was more important.

One possible objection to the analysis might be that listeners were biased by the inclusion of extrema in the random presentation of F0s. Therefore, in order to rule out the possibility that ranking of F0 extrema was accurate but ranking of more central F0 values inaccurate, linear regressions on individually defined tessitures were rerun with the top and bottom F0 for each speaker excluded. This test produced slightly lower correlation coefficients (inverse Fisher mean z of 0.609, down from 0.721), but, crucially, the correlation was still positive and significant at $p < 0.001$, even though n was 25% smaller. This test confirms the finding that listeners' successful ranking of pitch did not depend upon identifying peculiarities associated with the extreme high or low F0 values for each speaker.

The fact that listeners are able to locate within a tessiture steady, isolated pitches produced by unfamiliar voices implies some degree of speaker normalization. Although the present results strongly support our hypothesis that parsing of the F0 location does not require context or familiarization, they do not allow us to distinguish between possible models of F0 normalization. Listeners might build models of individual speakers or they might fit individual pitch ranges to previously built prototypes of possible vocal ranges. It is known that factors extrinsic to the signal are implicated in normalization-factors such as the overall formant range (Ladefoged and Broadbent, 1957) or the acoustics of inspi-

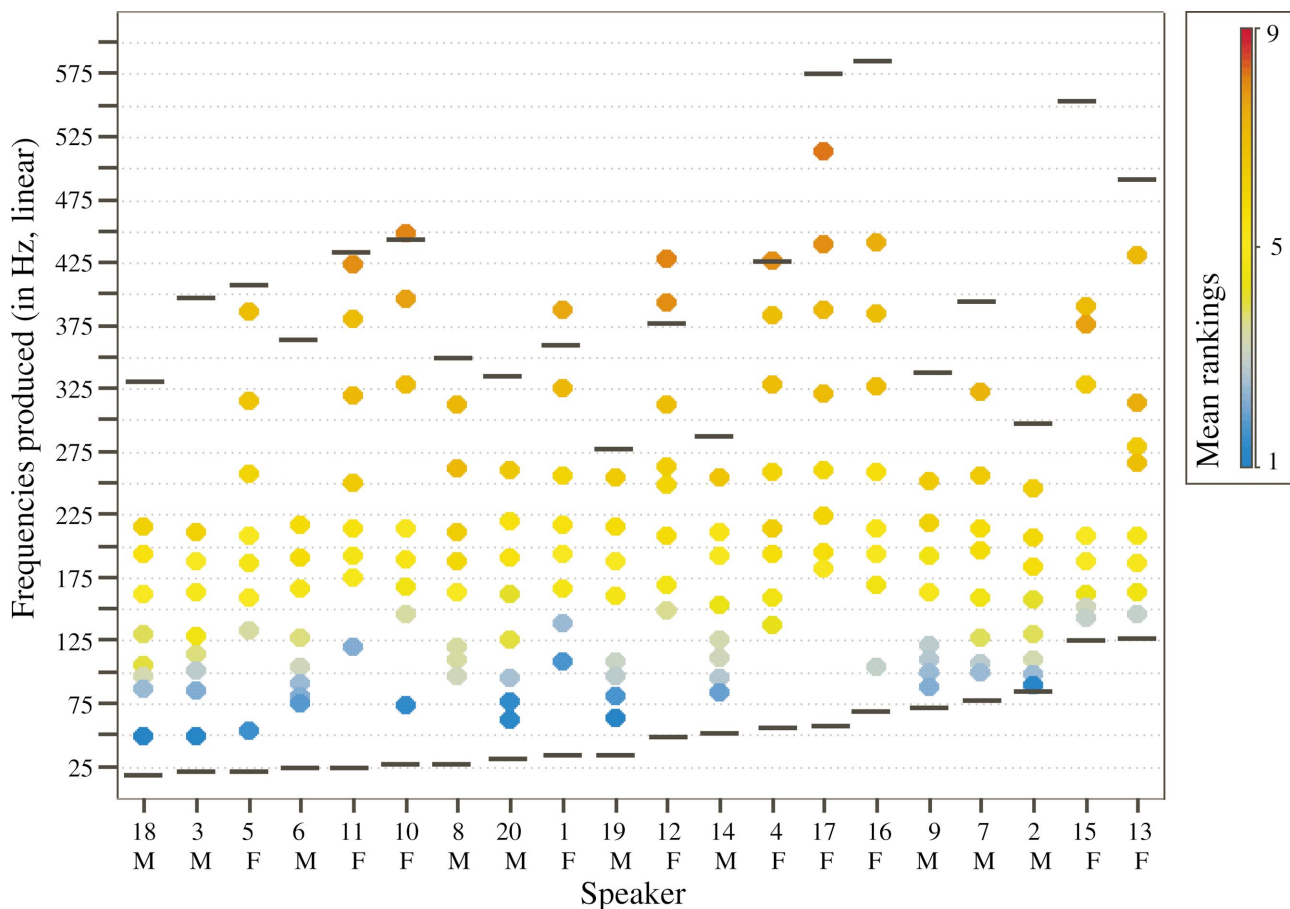


FIG. 4. Color-scaled mean rankings (12 listeners) plotted for each of eight sustained frequencies produced by each of 20 speakers. Data are ordered by bottom of glissando-based tessiture. The sex of the speaker is given below the speaker number.

ration, for example (Whalen and Sheffert, 1997). If signal-extrinsic aspects of speech are used in speaker normalization, it stands to reason that signal-intrinsic aspects of speech might also be used as listeners familiarize themselves with the voices of individual vocal tracts or vocal tract types. Voice quality is one such a signal-intrinsic feature.

Extremely low F0 within a speaker's range is often associated with a glottal fry or creaky voice (Ladefoged, 1971; Laver, 1980). Similarly, breathy voice tends to occur with low F0s (Fairbanks, 1960). Thus it may be that creak and breathiness indicate to the listener that the lower portion of the tessiture is being used. At the high end, the mode of vibration of the vocal folds changes dramatically between modal and falsetto voice (Hollien *et al.*, 1971; Laver, 1980). Minimally, therefore, it is likely that listeners can use information about voice quality directly to gauge the location of a steady pitch at the very least at the top and bottom of a speaker-specific range. However, in the present study, listeners appeared to base their rankings on signal-intrinsic (perhaps voice quality) information at or near the extrema as well as throughout the tessitures.

The proposal that listeners rely at least somewhat on voice quality as they locate pitch within a speaker-specific range does not lead naturally to a relational theory of F0 normalization in which pitch is located relative to other pitches via a Fraction-of-Range comparison or Z-Score-around-Mean comparison (e.g., Rose, 1987). In fact, one finding in the present investigation suggests the possibility of

a more direct route to F0 normalization: Listeners performed well throughout a speaker's range, even though F0s were presented randomly. If listeners use information on voice quality intrinsic to the signal to aid in predicting pitch location within the range, no model of speakers is needed, personal or generic. If listeners parse voice quality information and relate F0 information to it, normalization may follow a more direct route. The present findings do not argue conclusively in favor of voice quality as the basis of parsing, but they show clearly that listeners need not depend on sampling the F0s of a speaker before making range judgments. The slight improvement in the correlations when the F0s were coded for the full range of F0s in the experiment suggests that listeners might be calibrating their responses to their experience with typical F0 ranges that they have heard throughout their lives, with or without reference to the intricacies of voice quality changes at different parts of the F0 range. Since our speakers seem to be typical in this regard, such a strategy seems to be a reasonable one. Further tests, in which speakers with less overlapping ranges are selected (if this is possible), would be necessary to sort this out.

In addition to voice quality, there are many sources of information about pitch range available to listeners. However listeners use that information, we can be fairly certain that they do not perform normalization simply by linking speaker pitch location to the perceived size of the speaker. While it appears that vocal tract length in humans depends on skull size, and while it has been shown that vocal tract length, in

turn, directly affects the averaged distance between successive formant frequencies when the cross-sectional area is held constant (Fant, 1960; Fitch, 1997), neither average nor basal pitch (e.g., Cooper and Yanagihara, 1971) correlates directly with adult human vocal tract size. The perhaps common but false supposition of a body size-pitch connection may follow from our more generally correct intuition that large objects resonate at low frequencies, combined with a casual confusion between source and filter properties. Indeed, certain species of birds and mammals “falsely advertise” their body size with disproportionately low calls produced primarily by enlarging the vocal tract (Fitch, 1999; Fitch and Reby, 2001), but actual correlations between body size and the length and mass of the vocal folds have been found to be small in adult humans (e.g., Lass and Brown, 1978; Künzel, 1989) and in other vertebrates more generally (see Fitch, 1997 for a review). In fact, the larynx develops independently of the head and body, and is even hypertrophied in post-pubescent human males (Hollien, 1960). It seems that there is no direct link between larynx size (and thus average or basal F0) and vocal tract length or body size, so pitch normalization cannot proceed from “on-line” (probabilistic) estimation of vocal tract length.

Although F0 does not directly correlate with vocal tract length (or formant distance), a speaker’s sex can be determined reliably from either long (Lass *et al.*, 1978) or short (Ingemann, 1968) segments of typical speech, and automatic techniques are similarly successful with just vowels (Childers and Wu, 1991). It would appear that information regarding speaker sex may be present in what we might generally refer to as source or hybrid source-filter properties. It has been claimed, for example, that information about sex is available to the listener in the absolute value in Hz of some aspect of F0, usually mean or median F0. Peterson and Barney (1952) report male/female differences of around 100 Hz, as does van Bezooeyen (1984). Indeed, Bachorowski and Owren (1999) found that sex was classified correctly 92.4% of the time by discriminant analysis on the basis of formant frequency cues to vocal tract length for 2500 tokens of one vowel produced by 125 subjects. The success of a multivariate statistical technique in classifying speaker sex on the basis of vocal tract length cues alone does not necessarily mean that listeners judge speaker sex in the same way, or that listeners use judgments of speaker sex to normalize pitch at a gross level, but the possibility remains. It has also been claimed that information about sex may be available in spectral tilt. For example, Hanson and Chuang (1999) found the ratio of the amplitude of H1 to that of F3 to be 9.6 dB lower for male than for female speakers, with lower mean values and less variation between male speakers for F1 bandwidth, open quotient, and other measures. However, it is not known whether information about speaker sex is available to the listener in short segments throughout the tessiture. It may be that F0s at the high or low end can be misleading about the sex of the talker, with low values perhaps sounding more typically male and high values, female. This question will be addressed in future research. For now, it is reasonable to assume that the sex of the speaker for at least some of the stimuli could have been perceived in this experiment.

Given the listener’s potential for identifying the sex of the speaker, we might anticipate the following interpretation of our results, namely that listeners rank F0s based on stored templates representing population tessitures for adult male and adult female speakers, respectively. Indeed, correlations based on F0 locations within sex-based population ranges were almost as high as those based on global (sex-independent) ranges. Figure 4 shows evidence for both specific and global trends. If listeners were simply fitting judgments to absolute high- or low-frequency values, the colors of the dots would be expected to appear in straight horizontal bands uniformly throughout the chart; they do not, especially at the high and low ends. Yellow and near-yellow bands in the mid-tessiture range do cluster roughly in the 125–225 Hz range, but so do tessiture midpoints. In other words, where the yellow bands seem too high or too low relative to steady F0s produced by a speaker, they tend to fall roughly on center relative to speaker-specific tessitures. Furthermore, if listeners were simply fitting judgments to male and female (adult) population tessitures, the colors would be expected to band within sex, but plotting them this way (not shown directly) does not reveal any clear difference. In most cases where the lowest F0 within a speaker-specific tessiture does not receive a blue ranking or the highest F0, a red ranking, it is because of gaps in the stimuli. That is, there are steady F0s that, in an ideal world, could have been produced, but were not produced by some of our speakers. Major gaps appear on the low end for speakers 6 and 8 among the males and 4, 11, 12, and 17 among the females, and, on the high end, for speakers 2, 3, 6, 9, 18, and 20 among the males and 15 and 16 among the females. If the reader imagines dark blue and dark red dots in these spaces, respectively, a sex-independent pattern is enhanced. The dark dots simply follow tessiture boundary marks; they do not form bands of a single color family across all speakers of a particular sex. The lack of very high rankings may also be due to our intentional exclusion of falsetto from the stimuli. To the extent that falsetto is part of the speech range, listeners would be correct in not assigning a score of 9 to the highest F0s presented.

Where the F0 literature contains multiple studies of mean or median fundamental frequency as noted above, we focused on the entire range, and found a considerable overlap in tessiture between male and female speakers both in range and location of range. The overlap is especially apparent at the low end, where there is very little in the way of a consistent sex-based F0 difference in the glissando base. What we do see are more gaps for the females, who seem to have had a difficult time sustaining low F0s to which they can indeed glide down in a spoken voice register. At the high end, the highest glissando breaks are achieved by the females, and there is less overlap between sexes. There is some overlap, however, but it is again accompanied by (complementary) high-end gaps in sustained-but-possible F0s by the males. Had listeners here relied on a first-pass discrimination of sex to provide two templates for comparison, they would have misranked far more of these tokens. This pattern is consistent with the evidence for the use of both individual and global ranges.

Whether or not voice quality alone routinely provides

listeners with sufficient information for F0 normalization, it is clear that listeners are sensitive to the location of pitch within a speaker's range, even for isolated utterances. It will take further experimentation to isolate the features that listeners are sensitive to and to explain why listeners are good but not perfect at the task. For example, it would prove useful to test whether varying synthetic voice quality across a range of F0s confounds the listeners' attempts to track F0 values. Another avenue of research suggested by the present study would involve a replication in which the experimenters selected speakers with tessitures located high and low in absolute frequency and with less overlap than seen among the present speakers, if such speakers can be found.

ACKNOWLEDGMENTS

We thank Jeff Weihing for assistance with MATLAB programming and Brook Swainson and Paul De Decker for their assistance with stimulus extraction. We also thank Carol Fowler, Gordon Ramsay, and Yi Xu for their comments and Bob Ladd for a helpful discussion of the data reported here. During the review process, this paper also benefited from a number of insights offered by the Peter Assmann, Tecumseh Fitch and an anonymous reviewer. Any shortcomings in the present research remain our own. This work was supported by NIH Grant No. DC-02717 and NIH Grant No. DC-03782 to Haskins Laboratories and by an equipment loan from the Institute for Voice and Speech.

Bachorowski, J.-A., and Owren, M. J. (1999). "Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech," *J. Acoust. Soc. Am.* **106**, 1054–1063.

Childers, D. G., and Wu, K. (1991). "Gender recognition from speech. Part II: Fine analysis," *J. Acoust. Soc. Am.* **90**, 1841–1856.

Cooper, M., and Yanagihara, N. (1971). "A study of the basal pitch level variations found in the normal speaking voices of males and females," *J. Commun. Disord.* **3**, 261–266.

Di Paolo, M., and Faber, A. (1990). "Phonation differences and the phonetic content of the tense-lax contrast in Utah English," *Lang. Var. Change* **2**, 155–204.

Fairbanks, G. (1960). *Voice and Articulation Drillbook* (Harper and Row, New York).

Fant, G. (1960). *Acoustic Theory of Speech Production* (Mouton, The Hague).

Fant, G., Liljencrants, J., and Lin, Q. (1985). "A four-parameter model of glottal flow," *Speech Transmission Laboratory—Quarterly Progress and Status Report*, Royal Institute of Technology, Stockholm 4, pp. 1–13.

Fellowes, J. M., Remez, R. E., and Rubin, P. E. (1997). "Perceiving the sex and identity of a talker without natural vocal timbre," *Percept. Psychophys.* **59**, 839–849.

Fisher, R. A. (1915). "Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population," *Biometrika* **10**, 507–521.

Fitch, W. T. (1997). "Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques," *J. Acoust. Soc. Am.* **102**, 1213–1222.

Fitch, W. T. (1999). "Acoustic exaggeration of size in birds via tracheal elongation: Comparative and theoretical analyses," *J. Zool.* **248**, 31–48.

Fitch, W. T., and Reby, D. (2001). "The descended larynx is not uniquely human," *Proc. R. Soc. London, Ser. B* **268**, 1669–1675.

Gelfer, M. P. (1989). "Stability in phonational frequency range," *J. Commun. Disord.* **22**, 181–192.

Gobl, C., and Ní Chasaide, A. (1992). "Acoustic characteristics of voice quality," *Speech Commun.* **11**, 481–490.

Goldinger, S. D. (1996). "Words and voices: Episodic traces in spoken word identification and recognition memory," *J. Exp. Psychol. Learn. Mem. Cogn.* **22**, 1166–1183.

Hanson, H. M., and Chuang, E. S. (1999). "Glottal characteristics of male speakers: Acoustic correlates and comparison with female data," *J. Acoust. Soc. Am.* **106**, 1064–1077.

Heylen, L., Wuyts, F. L., Mertens, F., De Bodt, M., and Van de Heyning, P. H. (2002). "Normative voice range profiles of male and female professional voice users," *J. Voice* **16**, 1–7.

Hollien, H. (1960). "Some laryngeal correlates of vocal pitch," *J. Speech Hear. Res.* **3**, 52–58.

Hollien, H., Brown, W. S., and Hollien, K. (1971). "Vocal fold length associated with modal, falsetto, and varying intensity phonation," *Folia Phoniatr.* **23**, 66–78.

Ingemann, F. (1968). "Identification of the speaker's sex from voiceless fricatives," *J. Acoust. Soc. Am.* **44**, 1142–1144.

Johnson, K., and Mullennix, J. W. (ed.). (1997). *Talker Variability in Speech Processing* (Academic, San Diego, CA).

Klatt, D. H., and Klatt, L. C. (1990). "Analysis, synthesis, and perception of voice quality variations among female and male talkers," *J. Acoust. Soc. Am.* **87**, 820–857.

Künzel, H. J. (1989). "How well does average fundamental frequency correlate with speaker height and weight?," *Phonetica* **46**, 117–125.

Ladefoged, P. (1971). *Elements of Acoustic Phonetics* (University of Chicago Press, Chicago).

Ladefoged, P., and Broadbent, D. E. (1957). "Information conveyed by vowels," *J. Acoust. Soc. Am.* **29**, 98–104.

Lass, N. J., and Brown, W. S. (1978). "A correlational study of speakers' heights, weights, body surface areas, and speaking fundamental frequencies," *J. Acoust. Soc. Am.* **63**, 1218–1220.

Lass, N. J., Mertz, P. J., and Kimmel, K. L. (1978). "Effect of temporal speech alterations on speaker race and sex identifications," *Lang Speech* **21**, 279–290.

Laver, J. (1980). *The Phonetic Description of Voice Quality* (Cambridge University Press, Cambridge).

Leather, J. (1983). "Speaker normalization in perception of lexical tone," *J. Phonetics* **11**, 373–382.

Moore, C. B., and Jongman, A. (1997). "Speaker normalization in the perception of Mandarin Chinese tones," *J. Acoust. Soc. Am.* **102**, 1864–1877.

Nygaard, L. C., and Pisoni, D. B. (1998). "Talker-specific learning in speech perception," *Percept. Psychophys.* **60**, 355–376.

Palmeri, T. J., Goldinger, S. D., and Pisoni, D. B. (1993). "Episodic encoding of voice attributes and recognition memory for spoken words," *J. Exp. Psychol. Learn. Mem. Cogn.* **19**, 309–328.

Peterson, G. E., and Barney, H. L. (1952). "Control methods used in a study of the vowels," *J. Acoust. Soc. Am.* **24**, 175–184.

Reich, A. R., Frederickson, R. R., Mason, J. A., and Schlauch, R. S. (1990). "Methodological variables affecting phonational frequency range in adults," *J. Spinal Disord.* **55**, 124–131.

Remez, R. E., Fellowes, J. M., and Rubin, P. E. (1997). "Talker identification based on phonetic information," *J. Exp. Psychol. Hum. Percept. Perform.* **23**, 651–666.

Rose, P. (1987). "Considerations in the normalisation of the fundamental frequency of linguistic tone," *Speech Commun.* **6**, 343–351.

Schutte, H. K., and Seidner, W. (1983). "Recommendation by the Union of European Phoniaticians (UEP): Standardizing voice area measurement/phonetography," *Folia Phoniatr.* **35**, 286–288.

Sheffert, S. M., and Fowler, C. A. (1995). "The effects of voice and visible speaker change on memory for spoken words," *J. Mem. Lang.* **34**, 665–685.

Swerts, M., and Veldhuis, R. (2001). "The effect of speech melody on voice quality," *Speech Commun.* **33**, 297–303.

Titze, I. (1995). "Speaking vowels versus singing vowels," *J. Singing* **52**, 41–42.

van Bezoooyen, R. (1984). *Characteristics and Recognizability of Vocal Expressions of Emotion* (Foris Publications, Dordrecht).

Whalen, D. H., and Sheffert, S. M. (1997). "Normalization of vowels by breath sounds," in *Talker Variability in Speech Processing*, edited by K. Johnson and J. W. Mullennix (Academic Press, San Diego), pp. 133–144.

Wong, P. C. M., and Diehl, R. L. (2003). "Perceptual normalization for inter- and intratalker variation in Cantonese level tones," *J. Speech Lang. Hear. Res.* **46**, 413–421.

Zraick, R. I., Nelson, J. L., Montague, J. C., and Monoson, P. K. (2000). "The effect of task on determination of maximum phonational frequency range," *J. Voice* **14**, 154–160.

Perception of aperiodicity in pathological voice

Jody Kreiman^{a)} and Bruce R. Gerratt^{b)}

Division of Head and Neck Surgery, UCLA School of Medicine, 31-24 Rehab Center,
Los Angeles, California 90095-1794

(Received 20 December 2003; revised 20 December 2004; accepted 22 December 2004)

Although jitter, shimmer, and noise acoustically characterize all voice signals, their perceptual importance in naturally produced pathological voices has not been established psychoacoustically. To determine the role of these attributes in the perception of vocal quality, listeners were asked to adjust levels of jitter, shimmer, and the noise-to-signal ratio in a speech synthesizer, so that synthetic voices matched naturally produced tokens. Results showed that, although listeners agreed well in their judgments of the noise-to-signal ratio, they did not agree with one another in their chosen settings for jitter and shimmer. Noise-dependent differences in listeners' ability to detect changes in amounts of jitter and shimmer implicate both listener insensitivity and inability to isolate jitter and shimmer as separate dimensions in the overall pattern of aperiodicity in a voice as causes of this poor agreement. These results suggest that jitter and shimmer are not useful as independent indices of perceived vocal quality, apart from their acoustic contributions to the overall pattern of spectrally shaped noise in a voice. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1858351]

PACS numbers: 43.71.Bp, 43.71.Gv [PFA]

Pages: 2201–2211

I. INTRODUCTION

Jitter, shimmer, and noise-to-signal ratios (NSRs) are the cornerstones of acoustic measurement of voice signals, and are often regarded as indices of the perceived quality of both normal and pathological voices. In general, applications of acoustic measures to assess vocal quality derive their validity from the relevance of specific acoustic properties of the signal to auditory perceptions of voice. Researchers typically use correlation or regression techniques to demonstrate the extent to which such measures explain or predict listeners' scalar quality judgments. However, observed associations between acoustic and perceptual measures have varied considerably across studies. For example, correlations between measures of jitter and ratings of both breathiness and roughness have ranged from 0 to about 0.7 (for review see Heiberger and Horii, 1982; Kreiman and Gerratt, 2000). Multidimensional scaling studies (which examine the perceptual dimensions that underlie perceived vocal similarity) have also provided inconsistent results with respect to the role that jitter, shimmer, and the NSR play in determining quality (Kreiman *et al.*, 1990; Kreiman and Gerratt, 1996). Such variability in results has undermined confidence in both the acoustic measures and their perceptual interpretation.

Although hundreds of studies describing, evaluating, and applying measures of noise and acoustic signal perturbation have been published (Buder, 2000), the perceptual salience of these attributes remains poorly understood. A discrepancy exists between the results of early synthesis studies and findings from later investigations examining this association in naturally produced voices (Heiberger and Horii, 1982). Synthesis studies (Wendahl, 1963, 1966a, b; Coleman and Wendahl, 1967) used sawtooth waves with added jitter (± 1 –50 Hz around a mean f_0 of 100 or 200 Hz) or shimmer

(alternate periods reduced in amplitude by 1–6 dB). Near-perfect correlations were observed between the amount of jitter or shimmer and judgments of relative roughness for these nonspeech stimuli. More recently, Hillenbrand (1988) used synthetic vowels to examine the univariate relationships between jitter, shimmer, and noise and ratings of breathiness and roughness. He reported that roughness ratings increased with levels of jitter and shimmer up to about 2% jitter and 2-dB shimmer, after which increasing jitter and shimmer levels did not increase perceived roughness. Similarly, as the NSR increased, so did breathiness ratings.

Several limitations are inherent in these synthesis studies. The synthesis techniques employed in the earliest studies (Wendahl, 1963, 1966a, b) did not include vocal-tract filtering, and used a highly unnatural sawtooth source along with levels of jitter and shimmer (up to $\pm 50\%$) that greatly exceed those usually associated with the human voice (e.g., Andrianopoulos *et al.*, 2001; Munoz *et al.*, 2003). Stimuli lacked pitch contours or formant structure, and varied in only one acoustic dimension at a time. Further, a relatively small number of stimuli differed in relatively large acoustic steps, making it easy for listeners to perform the paired comparison task reliably (Heiberger and Horii, 1982). These factors could account for the high correlations with perceived vocal roughness, and limit the extent to which early studies explain how listeners hear naturally produced voice signals. (Hillenbrand, 1988, also commented that his stimuli sounded somewhat unnatural.) Further, previous studies assessing the role of noise and perturbation in determining voice quality have always assessed quality in terms of specific scales like breathiness, roughness, or hoarseness. However, the reliability of such scales has been repeatedly questioned, and their validity as measures of quality is also questionable (e.g., Jensen, 1965; Kreiman, Gerratt, and Berke 1994; see Kreiman *et al.*, 2005, for review).

A final limitation of previous synthesis studies is the fact

^{a)}Electronic mail: jkreiman@ucla.edu

^{b)}Electronic mail: bgerratt@ucla.edu

that jitter, shimmer, and noise have been manipulated independently of one another, and the perceptual interactions among these three variables have not been investigated. Acoustically, these attributes are obviously related. For example, adding aspiration noise to a signal adds jitter and shimmer; adding jitter also adds shimmer as changes in period length move harmonics toward and away from vocal-tract resonances; and adding jitter and shimmer produces measurable changes in the NSR (Hillenbrand, 1987). Naturally produced voice signals include all these attributes, and separating them analytically has proven difficult (Hillenbrand, 1987; Fukazawa *et al.*, 1988; de Krom, 1993; Titze, 1995; Michaelis *et al.*, 1997; Murphy, 1999).

In contrast to synthesis studies, correlational investigations of jitter, shimmer, and noise perception in naturally produced voices have produced highly variable results. Many low or nonsignificant correlations have been reported (e.g., Martin *et al.*, 1995; de Krom, 1995; Deal and Emanuel, 1978; Prosek *et al.*, 1987). However, correlational approaches are of limited use in resolving the issues surrounding the perceptual importance of vocal aperiodicity, because they cannot provide evidence about cause and effect. Lack of information about how (or whether) an acoustic variable evokes perception of a particular vocal quality leaves researchers and clinicians to make assumptions without evidence about the validity and utility of such measures. For example, it is impossible to determine if statistically significant changes in some parameter are actually perceptually meaningful or not in the absence of a demonstrated cause and effect relationship between an acoustic variable and a perceptual outcome.

In this study, we used a method of adjustment task to examine the perceptual significance of jitter, shimmer, and the NSR. Listeners were asked to manipulate jitter, shimmer, and NSR levels in high-quality synthetic voice stimuli, until the synthetic voices matched naturally produced target voices. This approach avoids the use of verbal rating scales, because listeners compare the stimuli directly in terms of their overall similarity. Further, in the method of adjustment task, subjects manipulate acoustic variables directly, so the association between signal and percept need not be inferred from correlation. Finally, in this task the perceptual importance of acoustic variables is evaluated in the same complex, multidimensional context in which the variables naturally occur. In particular, this study examined the perceptual interactions of jitter and shimmer with noise. A previous study (Gerratt and Kreiman, 2001) found good agreement among listeners for the NSR. It is possible that, because noise comprises jitter and shimmer acoustically, listeners agree about overall levels of noise, but not about the levels of jitter and shimmer present in a voice. On the other hand, if measures of jitter and shimmer are valid indices of perceived vocal quality, then listeners should agree well in the levels of these variables they select in the method of adjustment task.

II. EXPERIMENT 1

A. Rationale

To examine the contributions of jitter, shimmer, and noise to voice quality, listeners in this experiment were asked

to adjust the levels of jitter, shimmer, and/or noise present in a synthetic stimulus until that stimulus matched a naturally produced voice sample as closely as possible. We then examined the extent to which listeners agreed about how much of each variable was necessary for the synthetic and natural voice samples to sound the same, consistent with the ANSI definition of sound quality as “that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar” (ANSI Standard S1.1.12.9, p. 45, 1960). If jitter, shimmer, and noise contribute independently to the perceived quality of a voice, then listeners should agree well in the values they choose when making their adjustments.

B. Method

1. Voice samples

The voices of 20 speakers with voice disorders (10 males and 10 females) were selected from a large library of samples recorded under identical conditions as part of a clinical voice evaluation. Selection was random except that samples with period doubling or biphonation were excluded because jitter and shimmer are undefined for such signals (e.g., Titze, 1995). Speakers ranged in age from 22 to 78 years (mean=49.8 years; sd=17.2 years) and represented a variety of primary diagnoses, including mass lesions of the vocal folds (7), glottal incompetence (6), chronic laryngitis (4), adductory spasmodic dysphonia (2), and Parkinson disease (1). Severity of vocal deviation was rated by consensus vote of the authors. Ratings ranged from 2–5 on a 5-point scale (where 1=normal quality), and averaged 3.55 (sd =0.97).

During the voice evaluation, speakers sustained the vowel /a/ as steadily as possible. Voice signals were transduced with a 1-in. Bruel & Kjaer condenser microphone held a constant 5 cm off axis. They were then low-pass filtered at 8 kHz and directly digitized at 20 kHz. A 1-s segment was excerpted from the middle of these productions, antialias filtered, and downsampled to 10 kHz for further analysis.

2. Analysis and synthesis methods

Analysis techniques have been described in detail elsewhere (Gerratt and Kreiman, 2001; Gabelman, 2003). Briefly, formant frequencies and bandwidths were estimated using linear predictive coding analysis with a window of 25.6 ms (increased to 51.2 ms when stimulus f_0 was near or below 100 Hz). The glottal volume velocity derivative was estimated by interactively inverse filtering a single glottal pulse from the microphone recordings, using the method described by Javkin *et al.* (1987). The output of the inverse filter was least-squares fit with a modified Liljencrants–Fant (LF) source model (Fant *et al.*, 1985), and the parameters of the best-fitting LF model were used to specify the harmonic component of the voice source in the synthesizer. The shape of the inharmonic part of the voice source (the noise spectrum) was estimated by applying cepstral comb filtering to remove periodic source components following the method described by de Krom (1993). The residual was then inverse filtered to remove the effects of vocal-tract resonances, leav-

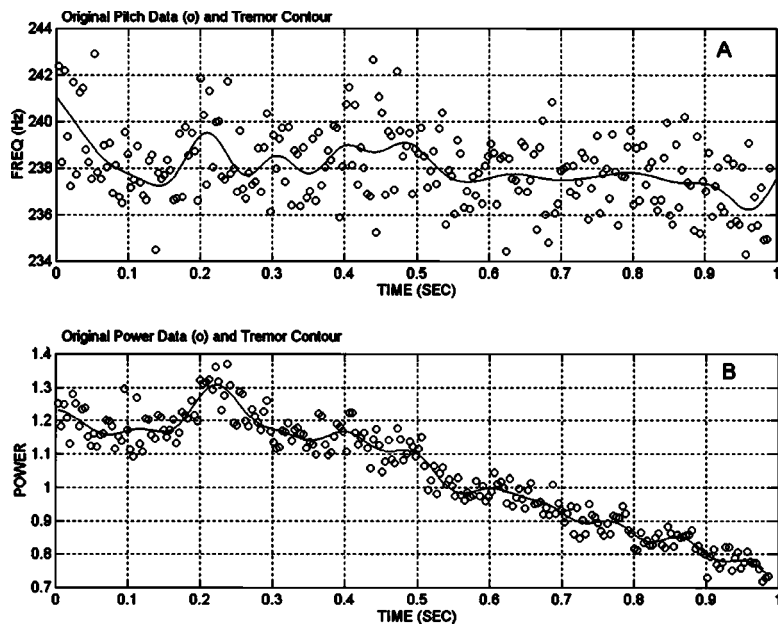


FIG. 1. Modeling of jitter, shimmer, and low-frequency modulations of frequency and power (tremors). (A) Output of f_0 analysis algorithm. Solid line shows the low-frequency pitch track, and open circles show deviations about this contour (jitter). (B) Output of power analysis algorithm. Solid line shows the low-frequency power track, and open circles show deviations about this contour (shimmer). Units for power are arbitrary.

ing the inharmonic part of the source, which was fitted with a 25-segment piecewise linear approximation. Finally, f_0 was tracked pulse by pulse on the time-domain waveform by an automatic algorithm. Location of cycle boundaries was verified by the first author. Tremor rates were estimated by visual inspection of smoothed plots of f_0 versus time (Kreiman *et al.*, 2003). Estimated tremor rates averaged 6.45 Hz, with a range of 2 to 10 Hz and a median of 6 Hz.

Variability in f_0 was modeled as follows (Fig. 1). The estimated tremor rate served as a dividing line between slow frequency and amplitude modulations (tremors) on the one hand and fast variations (jitter and shimmer) on the other. To model tremor, the f_0 track was low-pass filtered with the cutoff value equal to the estimated tremor rate, and period lengths were altered to incorporate this slowly varying frequency track [Fig. 1(a)]. Low-frequency amplitude modulations were similarly generated for each voice [Fig. 1(b)], and the power of each cycle was modified to incorporate observed low-frequency modulations in amplitude.¹ Jitter was modeled by altering the duration of each cycle by an amount sampled from a high-pass filtered, normally distributed random sequence whose variance was determined by the desired level of jitter. The filter's cutoff frequency again equaled the estimated tremor value. Shimmer was similarly modeled by altering the power of each cycle of phonation, with power values sampled at random from a normal distribution of values whose variance corresponded to the desired amount of shimmer. (See Gabelman, 2003, for more detail.)

Stimuli were synthesized using a custom formant synthesizer implemented in MATLAB (MathWorks, Natick, MA; Gerratt and Kreiman, 2001).² The synthesizer sampling rate was fixed at 10 kHz. The following procedure was applied to overcome quantization limits on modeling f_0 . Source pulses were synthesized pulse by pulse. A single LF-modeled source pulse was upsampled from 10 to 40 kHz and used as a basis for each pulse in the time series. A plot of the desired f_0 values versus time was generated for the 1-s synthetic sample, taking into account the tremor contour and requested

level of jitter. As the series of pulses was constructed, precise ending times for each LF pulse were tracked, and this curve was interpolated to find f_0 for the next pulse in the series. The basic LF pulse was then time warped (stretched or compressed) to obtain the appropriate period length, and sample points were calculated accordingly. A similar power versus time curve was constructed for the time series, reflecting amplitude tremor and shimmer, and this curve was interpolated to determine the power for each successive cycle in the time series. The overall effect is equivalent to digitizing an analog pulse train with pulses of the exact desired frequencies and amplitudes at the fixed 10-kHz sample rate.

A 100-tap finite impulse response filter was constructed for noise synthesis, with its shape corresponding to the 25-segment piecewise linear approximation fit to the inharmonic part of the voice source derived through comb filtering. A spectrally shaped noise time series was then created by passing white noise through this filter. The jittered, shimmered, and tremored LF pulse train was antialias filtered and down-sampled to 10 kHz, and then added to this noise time series to create a complete source time series. The source was filtered through the vocal-tract model to generate a preliminary version of the synthesized voice. Finally, formant frequencies, bandwidths, and LF parameters were adjusted to provide good perceptual matches to the target voices, and then held constant across experimental conditions.

3. Listening pretest

A listening pretest was used to verify the accuracy of the synthesis. Prior to this pretest, estimated values of the NSR, jitter, and shimmer were adjusted in the synthesizer by the first author as necessary to provide the best possible perceptual match to the natural voice samples, because measurement of jitter, shimmer, and the NSR is difficult and often inaccurate when phonation departs from periodicity (e.g., Titze, 1994; Bielamowicz *et al.*, 1996). Synthetic copies of each of the 20 natural voice samples were then generated,

using the methods described above and these perceptually adjusted levels of jitter, shimmer, and noise. Twenty listeners (UCLA students and staff; 20 to 53 years of age; mean age = 26.4 years; $sd=9.93$ years) heard pairs of voices. On half of the trials, a synthetic voice sample was paired with its natural counterpart, and on the other half stimuli were identical. Each pair was repeated 3 times, for a total of 120 trials/listener.

For each trial, listeners were asked to judge whether the two samples were the same or different, and to rate their confidence in their response on a 5-point scale ranging from “positive” to “wild guess.” Listeners were not allowed to replay the stimuli before responding. Order of voices in “different” pairs was randomized, and the stimulus pairs were rerandomized for each listener. Listeners were tested individually in a double-walled sound suite. To approximate normal listening conditions, stimuli were presented in free field at a comfortable constant listening level. Testing lasted approximately 15 min.

To provide a measure of the average discriminability of the synthetic and natural tokens, responses were pooled across listeners. Overall rates of correct and incorrect “same” responses (hits and false alarms) were calculated for each voice. Hit rates ranged across voices from 85%–98.3% (mean=91.8%; $sd=3.66\%$); false-alarm rates ranged from 51.7%–85% (mean=65.7%; $sd=11.18\%$). The probability of any correct response (“same” or “different”) ranged across voices from 54.2%–70% (mean=63.1%; $sd=5.02\%$).

Same/different responses for each voice were combined with confidence ratings to create a 10-point scale ranging from “positive voices are the same” to “positive voices are different.” For example, “same” responses with confidence equal to 1 (positive) were recoded as “1,” “same” responses with confidence equal to 5 (wild guess) were recoded as “5,” “different” responses with confidence equal to 5 were recoded as “6,” and “different” responses with confidence equal to 1 were recoded as “10.” Receiver operating characteristics (ROCs) consisting of 9 points each (10 recoded response categories minus 1) were constructed from these recoded data following the procedure described by Green and Swets (1966; see also MacMillan and Creelman, 1991). The area under the ROC for each voice was calculated, along with 99% confidence intervals around these values. In all cases, these confidence intervals included the chance value of 0.5. These data, combined with consistently high false-alarm rates, indicate that listeners were unable to consistently distinguish the synthetic copies from the natural samples. We conclude that the synthesis technique is able to model the quality of the natural target voice samples adequately for purposes of the following experiments.

4. Listening task

Seventy listeners participated in experiment 1. They ranged in age from 19–57 years (mean age=26.7 years; $sd=7.83$ years). All reported normal hearing. Listeners were unselected with respect to experience with voice disorders, and most were phonetically- and otolaryngologically naive.

Listeners were tested individually in a sound-treated room. Each listener participated in 20 trials, one for each

voice. Voices were presented to each listener in a unique random order, in free field through two loudspeakers at a constant comfortable listening level (to approximate normal listening conditions). At the beginning of each trial, listeners heard the natural voice sample paired with a copy of that voice synthesized without jitter, shimmer, or noise. They were told that one voice was naturally produced, and that the other was a synthetic copy. They were then asked to adjust the quality of the second voice by moving a sliding cursor or cursors on a computer monitor with a mouse, until the two voices matched on the target dimension or dimensions as closely as possible. These cursors controlled the jitter, shimmer, and NSR levels present in the synthetic voice. Depending on condition, listeners adjusted either one, two, or all three parameters in a given trial. Listeners heard a given voice in a single condition only. Which task a listener performed for a given voice was assigned at random, with the constraint that ten listeners performed each of the seven possible tasks (jitter only; shimmer only; noise only; jitter + shimmer; jitter+noise; shimmer+noise; or jitter+shimmer+noise) for each voice. Listeners were encouraged to play the natural and synthetic stimuli as often as they liked in any order, and could make as many adjustments as necessary to achieve the best possible match to the target voice.

Scale displays were 115 mm in length. The NSR scale ranged from -50 to 0 dB (noise-free to extremely noisy); the jitter scale ranged from 0% to 3% ; and the shimmer scale ranged from 0 to 2 dB. These values were chosen based on measurements of the test voices, on pilot tests, and on data from Hillenbrand (1988), who reported that increasing jitter beyond 2.5% and shimmer beyond 2 dB had little additional effect on perceived vocal roughness. Ranges of values for jitter, shimmer, and the NSR vary widely across previous studies, for both pathological and normal voices, due to differences in measurement techniques, instrumentation, and computational algorithms. In the present research, stimuli were frequency- and amplitude demodulated prior to perturbation analysis to remove the contributions of tremor from measured values. Values of these variables in the synthesizer were also calculated independently of the effects of tremor. For these reasons, the ranges used here for these variables are lower than those reported by some authors (see Buder, 2000, for review). Parameters not being manipulated in a given trial remained set in their “off” position (0% jitter, 0 -dB shimmer, -50 -dB NSR).

Prior to the experiment, the synthesizer was demonstrated and two practice items were presented, so that listeners could become thoroughly familiar with the task and with the sounds of different amounts of jitter, shimmer, and noise, individually and in combination. Practice continued until subjects were satisfied that they understood and could perform the task; it lasted about 15 min on average. The total test session lasted 1.5–2 h. Listeners were encouraged to take breaks during this period as necessary to maintain attention and motivation.

C. Results

Figure 2 shows the distributions of jitter, shimmer, and NSR responses, summed over voices and experimental con-

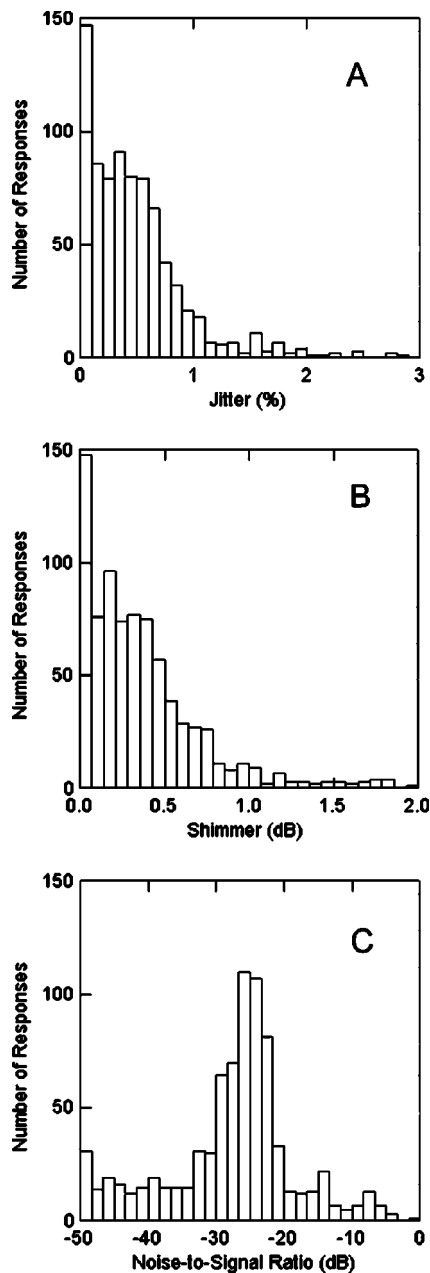


FIG. 2. Distribution of responses for jitter (A), shimmer (B), and the NSR (C), pooled across the 20 stimuli in the method-of-adjustment task in experiment 1.

ditions. Across conditions and voices, listeners used the full range for each of these scales. Responses for jitter and shimmer clustered toward the lower third of each scale, with listeners adding only small amounts of jitter and/or shimmer on most trials. In contrast, NSR responses were approximately normally distributed.

To measure response variability, we calculated the coefficient of variation (the standard deviation divided by the mean) for the jitter, shimmer, and NSR responses for each voice. Figure 3 shows the distribution of variation coefficients for each measure across the 20 stimulus voices. For each of the individual voices, variability of the jitter and shimmer responses exceeded that of the NSR responses. On average, jitter and shimmer responses were more than 5 times more variable than NSR responses, and in one case

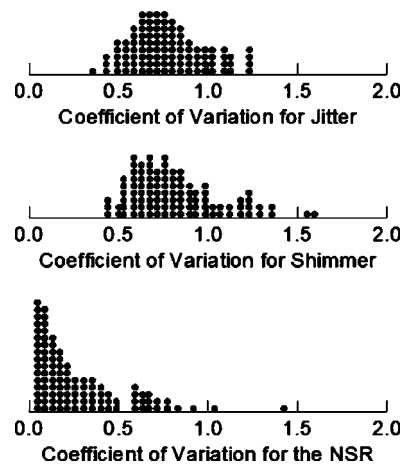


FIG. 3. Distribution of coefficient of variation values for the jitter, shimmer, and NSR responses in experiment 1, across experimental conditions and stimuli.

they were over 20 times more variable. These differences were significant across voices [$F(2,297) = 162.12$, $p < 0.01$; Bonferroni *post hoc* comparisons $p < 0.01$].

Jitter and shimmer responses varied significantly with the particular listening task. Listeners used significantly less jitter and shimmer when asked to adjust all three parameters at once than they did when matching jitter or shimmer alone [jitter: $F(3,796) = 4.32$, $p < 0.01$; shimmer: $F(3,796) = 4.14$, $p < 0.01$; Bonferroni *post hoc* comparisons $p < 0.01$]. In contrast, listeners always adjusted the NSR to similar levels, even when they were also adding jitter and/or shimmer to the voice, so no significant effect of response condition on NSR values was observed [$F(3,796) = 1.01$, n.s.]. However, coefficients of variation did not vary significantly by experimental task [jitter: $F(4,95) = 0.288$, $p > 0.01$; shimmer: $F(4,95) = 1.05$, $p > 0.01$; NSR: $F(4,95) = 0.525$, $p > 0.01$].

Variability in NSR responses *decreased* with increasing rated severity of vocal deviation. In other words, the worse the voice sounded, the better the listeners agreed in their NSR responses ($r = -0.64$, $p < 0.01$). No significant relationship between severity of deviation and response variability was observed for jitter or shimmer (jitter: $r = -0.23$, n.s.; shimmer: $r = -0.16$, n.s.). However, variability in jitter and shimmer responses did increase significantly with the NSR (jitter: $r = 0.68$, $p < 0.01$; shimmer: $r = 0.74$, $p < 0.01$). When the NSR was low, listeners' responses rarely exceeded 1% jitter or 1-dB shimmer. However, as the NSR increased, so did response variability, with some listeners adding little or no jitter or shimmer to the synthetic stimuli, and others adding large amounts. This contrasts sharply with NSR responses: The higher the NSR, the better listeners agreed in their responses ($r = -0.66$, $p < 0.01$).

To examine the relationship between response variability and the spectral shape of the harmonic part of the source (Gerratt and Kreiman, 2001), we calculated the difference in the amplitudes of the first two harmonics ($H1-H2$) from FFT spectra of the LF source pulses. This measure provides one index of the source spectral slope, independent of the influence of the vocal-tract transfer function: The larger

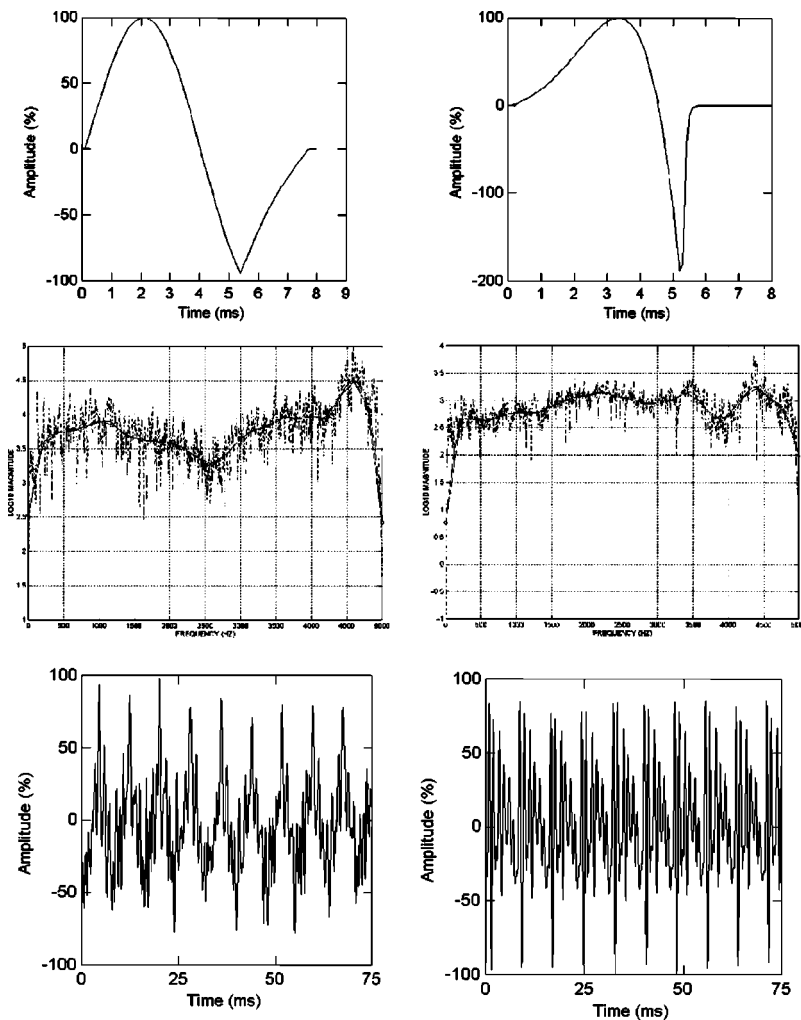


FIG. 4. Sources, noise spectra, and time series waveforms for representative stimuli used in experiment 2. The leftmost three panels represent a voice with a relatively sinusoidal source and a high noise-to-signal ratio (-6.9 dB). The rightmost three panels represent a voice with a nonsinusoidal source and a lower noise-to-signal ratio (-40.4 dB).

$H1-H2$, the more sinusoidal the source. Variability in NSR responses increased as $H1-H2$ decreased ($r = -0.57$, $p < 0.02$). Listeners agreed best in their NSR responses when the harmonic part of the voicing source was near-sinusoidal, as predicted by theory and previous research (Gerratt and Kreiman, 2001). That is, listeners were less sensitive to changes in the NSR when enough energy from the harmonic part of the voice source was also present to provide perceptually salient excitation in the high-frequency part of the voice spectrum. No such relationship was observed between harmonic source characteristics and agreement levels for jitter or shimmer (jitter: $r = 0.29$, n.s.; shimmer: $r = 0.28$, n.s.).

D. Discussion

These data suggest that listeners are relatively sensitive to the overall extent of aperiodicity in the signal, as measured by the NSR. Response variability was relatively low for NSR responses; responses were consistent across experimental conditions, and agreement increased with the amount of noise present. In contrast, the correspondence between jitter, shimmer, and perceived vocal quality appears far less precise. Responses varied widely when listeners were asked to match the amounts of jitter and shimmer present in natural voice signals. Jitter and shimmer responses varied significantly with experimental condition, and variability in re-

sponses increased with increasing aperiodicity. No effect of task on variability was observed, suggesting that listener difficulties were not related to the presence or absence of noise (and corresponding differences in the naturalness of the stimuli).

The dependence of jitter and shimmer responses on experimental condition suggests that listeners judge aperiodicity with respect to the overall amount of noise in the signal, rather than by decomposing aperiodicity into independent, separable aspects. This in turn suggests that listeners' relative insensitivity to the amounts of jitter and shimmer present in a voice signal may be due to their inability to isolate jitter and shimmer within a noisy voice signal. The increasing variability in jitter and shimmer responses as the NSR increased is consistent with this explanation. When the NSR was low, listeners chose (relatively) low values for all aspects of aperiodicity. As the NSR increased, however, variability in jitter and shimmer responses increased, suggesting that listeners as a group could not decide how much of the aperiodicity they heard was jitter or shimmer.

The following experiment tested this hypothesis. We synthesized stimuli with and without noise, and added jitter and shimmer in steps to those stimuli to create several series of voices. If listeners are able to separate the effects of jitter and shimmer from the overall noise pattern, then they should be equally able to detect differences between these stimuli

TABLE I. Stimulus characteristics, experiment 2.

Voice	Speaker sex	Source	NSR (dB)
1	F	-sine	-23.2
2	M	+sine	-6.9
3	F	+sine	-14.9
4	M	-sine	-40.4
5	M	+sine	-24.9
6	F	-sine	-29.2
7	F	+sine	-28.5
8	M	-sine	-29.5

whether noise is absent or present. If listeners are simply grossly insensitive to jitter and shimmer, then they should have difficulty detecting differences whether noise is present or absent. Finally, if listeners can easily distinguish among the variable jitter and shimmer responses from experiment 1, this would implicate difficulties with the multivariate matching task as the cause of the observed variability in responses.

III. EXPERIMENT 2

A. Method

1. Stimuli

Eight voices (four males and four females) were selected from the set of 20 studied in experiment 1. Because sensitivity to the level of noise in a voice may depend on the spectral shape of the harmonic part of the source (Gerratt and Kreiman, 2001), voices were chosen based on the shape of the LF-fitted source pulses (relatively sinusoidal/relatively non-sinusoidal) and on the NSR of the natural voice (relatively high/relatively low) (Fig. 4). Stimuli were chosen so that these parameters varied as orthogonally as possible, given the characteristics of the original voice samples. One male and one female voice were assigned to each source-by-NSR cell (Table I).

Five series of stimuli, each comprising five tokens, were synthesized for each of these eight voices. In two series, the amount of jitter increased across the five tokens in steps. One jitter series was synthesized with the NSR set at a constant value equal to the mean of the levels listeners selected for that voice in experiment 1. The second jitter series was synthesized with the NSR set at -50 dB (no perceptible noise present). Two series of shimmered stimuli were also created, one in which the NSR was set at -50 dB and one with the

NSR set at the mean of the values listeners selected in experiment 1. In both series, the amounts of shimmer present increased across the series in steps, as for the jittered stimuli. The fifth series of stimuli was synthesized with the amount of noise increasing in steps, without any additional jitter or shimmer. Because noise responses were independent of jitter and shimmer responses in experiment 1, additional stimulus series were not created varying noise levels in the contexts of average amounts of jitter and shimmer.

Series endpoints and step sizes were individually determined for each voice (Table II). Endpoints for the jitter series represented the maximum and minimum jitter responses observed for that voice in the condition in experiment 1 in which listeners adjusted jitter only, with noise and shimmer set at -50 and 0 dB, respectively. The three intermediate points in each series were evenly spaced (in acoustic units) between these extremes. Endpoints for the shimmer and noise series were similarly selected based on “shimmer only” and “noise only” response ranges from experiment 1. All other synthesis parameters were held constant for each voice at the values used in experiment 1.

Stimuli were 1 s in duration, and were synthesized with a sampling rate of 10 kHz, using the methods described for experiment 1. They were multiplied by 25-ms onset and offset ramps and scaled for equal peak amplitude prior to presentation to listeners.

2. Listeners and listening task

Eighteen listeners participated in this experiment. All reported normal hearing. For each series of stimuli for each voice, listeners heard all possible pairs of the five synthetic tokens in the series, plus an equal number of pairs where stimuli were the same, for a total of 800 trials/listener (8 voices \times 5 series/voice \times 10 comparisons/series, plus 400 “voices same” trials). They were asked to determine whether the stimuli were the same or different within a pair, and to rate their confidence in their response on a 5-point scale ranging from “positive” to “wild guess.”

Testing took place in a double-walled sound suite in two sessions, each lasting about 45 min. Stimuli were presented in free field at a comfortable constant listening level, and were rerandomized for each listener. Voices within a pair were separated by 350 ms. Listeners controlled the rate at which pairs were presented, but were not allowed to play a pair more than once before responding.

TABLE II. Characteristics of stimulus series for jitter, shimmer, and noise in experiment 2.

Voice	Step sizes for jitter (%)	Jitter continuum endpoints	Step sizes for shimmer (dB)	Shimmer continuum endpoints	Step sizes for NSR (dB)	NSR continuum endpoints
1	0.26	0.018, 1.05	0.18	0, 0.73	1.60	$-27.15, -20.80$
2	0.65	0.15, 2.75	0.39	0.18, 1.75	1.46	$-10.10, -4.25$
3	0.41	0, 1.63	0.45	0, 1.82	0.87	$-16.90, -13.40$
4	0.25	0, 0.98	0.10	0, 0.41	4.68	$-49.20, -30.45$
5	0.47	0.08, 1.94	0.30	0, 1.18	1.97	$-29.20, -21.30$
6	0.19	0, 0.74	0.15	0.04, 0.64	4.57	$-41.95, -23.65$
7	0.23	0, 0.90	0.24	0.09, 1.07	4.17	$-40.20, -23.50$
8	0.70	0, 2.80	0.29	0, 1.16	2.84	$-36.60, -25.25$

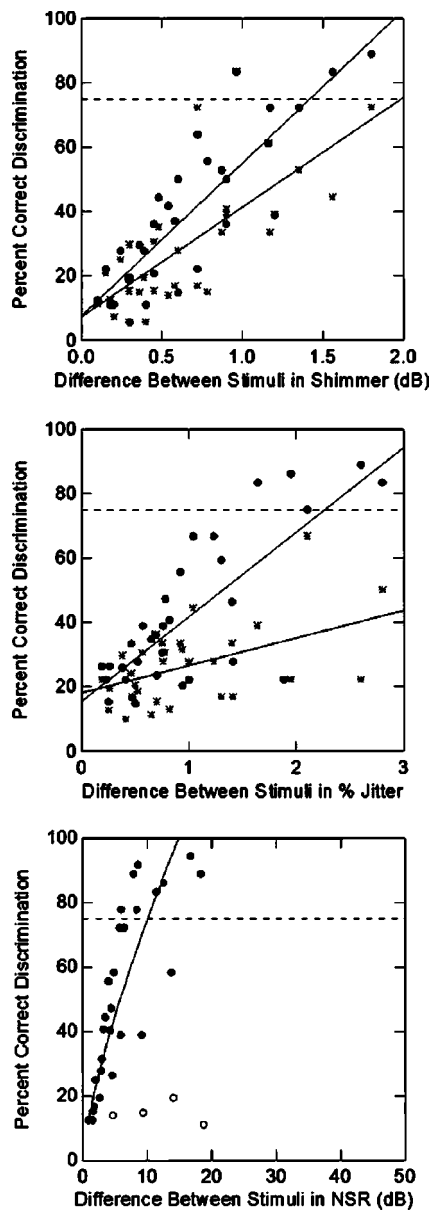


FIG. 5. The rate of correct “different” responses versus as a function of the acoustic distance between stimuli. For jitter and shimmer data, stimuli from series that included noise are plotted with stars, and stimuli from series synthesized without noise are plotted with filled circles. For noise data, open circles indicate a voice with a nonsinusoidal source and a high NSR. Regression lines through the two series of data in each panel are plotted with solid lines. The dashed line corresponds to 75% correct discrimination, as described in the text. Top panel: shimmer. Middle panel: jitter. Bottom panel: the noise-to-signal ratio.

3. Results and discussion

Same/different response data were pooled across listeners to estimate overall discrimination performance.³ For each voice and stimulus series, we calculated the percentage of correct “same” responses and the percentage of correct “different” responses. The rate of correct “same” responses was consistent across experimental conditions [$F(4,35) = 0.14, p > 0.05$], ranging from 78.9%–92.8% (mean = 86.5%, $sd = 3.70\%$). Discrimination accuracy (the rate of correct “different” responses) is shown as a function of the acoustic distance between stimuli within a pair in Fig. 5.

Each panel of this figure combines data from the eight individual stimulus voices, pooled across listeners.

Results for jitter and shimmer are shown in panels (A) and (B). In these panels, stimuli from series that included noise are plotted with stars, and stimuli from series synthesized without noise are plotted with filled circles. Regression lines through the two series of data in each panel are plotted with solid lines [Jitter: with noise, $F(1,30) = 9.12, p < 0.01$; no noise, $F(1,30) = 46.39, p < 0.01$. Shimmer: with noise, $F(1,30) = 32.33, p < 0.01$; no noise, $F(1,30) = 83.56, p < 0.01$]. Data in panels (A) and (B) show similar patterns. Discrimination accuracy was significantly better in both cases when noise was absent (filled circles) than when it was present (stars). For jitter, correct discrimination rates averaged 14.95% greater when noise was absent than when it was present [matched pairs $t(31) = 4.37, p < 0.01$]. For shimmer, discrimination accuracy averaged 9.03% more when noise was absent [matched pairs $t(31) = 3.74, p < 0.01$].

Difference limens were estimated by the point of intersection between the regression lines fit to the data and the point at which discrimination accuracy reached 75% (indicated by a dashed line in the figures; e.g., Marks and Algom, 1998). For jitter, averaged across the eight voices, listeners only reliably heard a difference between stimuli that differed by 2.27% or more in jitter when spectral noise was absent. When noise was present, listeners never reliably heard a difference between the stimuli used in this experiment, indicating that stimuli must differ by more than 3% jitter (the maximum range available) for the difference to be reliably perceptible. For shimmer, in the absence of spectral noise listeners could discriminate only among stimuli that differed by at least 1.42 dB in mean shimmer; in the context of spectral noise, a difference of 1.99 dB in mean shimmer (given a maximum range of 2.0 dB) was required for listeners to hear a difference between stimuli 75% of the time.⁴

The pattern is somewhat different for spectral noise [panel (C)]. One voice (number 4) with an extremely high NSR and a harmonic source that provided significant high-frequency excitation was an obvious outlier (plotted with open circles in the figure). When this voice is omitted from the analysis, the function relating discrimination accuracy to differences in noise levels appears curvilinear. This curve was interpolated using a power function rather than a linear function [$F(1,26) = 43.58, p < 0.01$], as shown in the figure, omitting the outlying case. For stimuli to be correctly discriminated 75% of the time required a difference in the NSR of 10.65 dB, a relatively low value given the 50-dB range of the scale. (Recall that listeners used the entire scale for all three measures, so observed differences in the magnitude of the estimated difference limens with respect to the length of the scales are not due to the length of the scales themselves.)

Experiment 2 examined three hypotheses regarding the causes of poor listener agreement in experiment 1. The present results are consistent with two of these hypotheses: Listeners are unable to isolate jitter and shimmer as separate components in the overall pattern of aperiodicity in a voice, and they are also insensitive overall to jitter and shimmer. Listeners were significantly better at discriminating levels of both jitter and shimmer when noise was absent than when it

was present, indicating that listeners do have difficulty decomposing the overall aperiodic component of a voice into independent constituent parts. However, estimated difference limens for jitter and shimmer are large overall, both in the presence and in the absence of spectral noise. These limens are also large relative to jitter and shimmer levels usually considered experimentally or clinically significant. For example, Karnell *et al.* (1995) reported that differences in the jitter measured by different analysis systems averaged 0.01%, while differences in shimmer averaged 0.085 dB. Bielamowicz *et al.* (1996) reported differences between systems in measured jitter of about 0.4%–0.5%; differences between systems in measured shimmer values were less than 0.1 dB. Linville (2000) reported that the voices of old and young men differed in jitter by about 0.6%. Shimmer values for old and young men differed by about 0.4 dB. Hanson *et al.* (1997) found decreases in jitter of about 0.03% after 6–9 weeks of treatment for laryngeal inflammation from gastroesophageal reflux, and claimed that these changes “document the changes in vocal quality with treatment for chronic laryngitis” (p. 284). Shimmer levels for patients with unilateral vocal-fold paralyses differed from those of control patients by about 1 dB (Hartl *et al.*, 2003); and reliable changes in jitter level during a histamine challenge test averaged about 1.5%, leading to the conclusion that “jitter is an objective and repeatable measurement of hoarseness” (Jones *et al.*, 2001, p. 29). Although probable variations in measurement techniques may limit comparisons among studies, estimated difference limens for jitter and shimmer in the present study far surpass these values, and also exceed the measurement precision usually required for such measures (Titze, 1995; Titze *et al.*, 1987).

Finally, the results do not support the hypothesis that listener problems in agreement observed in experiment 1 resulted from difficulties in performing the multivariate method-of-adjustment task. In our previous study using this method (Gerratt and Kreiman, 2001), listener agreement was substantially higher for NSR settings than it was for perceptual ratings of the noisiness of the same stimuli. However, in that study listeners were asked to adjust only a single synthesizer parameter. Similar levels of listener agreement were observed for NSR settings in the present study, when listeners were asked to adjust as many as three parameters simultaneously. Further, NSR ratings remained consistent, whether listeners were adjusting one, two, or three parameters simultaneously. The fact that listener agreement about NSR ratings remained consistently high across tasks suggests that listeners’ inconsistency with each other when matching jitter and shimmer levels was not due to the multivariate task, but instead was related to their auditory insensitivity to these two parameters.

IV. GENERAL DISCUSSION

Acoustic measures of voice derive their importance from the relevance of the acoustic signal to auditory perception of voice, association with some aspect of laryngeal physiology, or both (Catford, 1977). Jitter, shimmer, and noise have many physical sources and may arise at many stages in the speech production process (see, e.g., Titze, 1994, for review).

Because a given measurement value can reflect so many different causes, acoustic measures of aperiodicity are not diagnostically useful as indices of any particular physical state or physiological process. This leaves associations with voice quality as a possible motivation for measuring aperiodicity. This motivation also appears to fail in the cases of jitter and shimmer. Comparisons with published values are difficult because recording techniques, measurement procedures, and computational algorithms vary widely (see Buder, 2000, for review), and authors are not always specific about the methods applied. However, it appears that listeners are insensitive to the amounts of jitter and shimmer present in a voice sample within a range often treated as meaningful, as discussed above. We conclude that the associations between jitter, shimmer, and perceived voice quality are not sufficiently explanatory to justify continued reliance on jitter and shimmer as indices of voice quality.

The case is stronger for the perceptual relevance of NSR measures, because listeners agreed far better in their NSR responses than they did for jitter and shimmer, and the minimum reliably perceptible difference for the NSR was much smaller relative to the range of observations than it was for jitter and shimmer. NSR responses were independent of experimental task, and observed variability could be explained in part by the pattern of high-frequency excitation in a voice. These results indicate that listeners respond perceptually to changes in the NSR in consistent and principled ways, suggesting that the NSR is a significant and reliable determinant of vocal quality.

However, the fact that a measure is perceptually or psychologically important does not mean that researchers can ignore the limits of perceptual resolution on that scale when applying it. Although comparisons to values in the literature are again difficult (differences among computational methods for NSR measures being particularly inscrutable and vexing; see Buder, 2000), many authors’ claims about group differences or treatment effects rest on NSR differences that may be imperceptible. For example, Su *et al.* (2002) found statistically significant differences in pre- and postsurgical NSR values of 0.12 dB; Jotz *et al.* (2002) reported that each increase of 0.01 dB in the NSR doubled the risk of dysphonia in a sample of boys with and without vocal-fold lesions; and Niedzielska (2001) found differences in NSRs between control subjects and various diagnostic groups ranging from 2.7 to 14.2 dB. Further research examining the perceptual interactions between the harmonic and inharmonic parts of the voicing source should contribute to standardizing NSR measures so that they reflect vocal quality as accurately as possible. This will enhance our ability to apply NSR measures appropriately.

Finally, the present data extend our previous findings regarding the reliability of the method of adjustment task from expert listeners (Gerratt and Kreiman, 2001) to naive listeners. Reanalysis of data from our previous experiment, which used a different set of stimulus voices and expert listeners, produced a difference limen for the NSR (the point at which listeners achieved 75% correct discrimination, on average) of 13.4 dB, compared to the value of 10.65 dB for naive listeners in the present study. Expert and naive listeners

have previously been shown to differ significantly in the perceptual strategies they apply when rating pathological voices on traditional scales like breathiness or roughness (Kreiman *et al.*, 1990). However, the similarity in sensitivity levels for expert and naive listeners suggests that this method-of-adjustment task controls the effects of listener experience on perceptual responses, as previously predicted (Kreiman and Gerratt, 2000).

The results reported here may be surprising in light of the hundreds of papers published on these acoustic measures over the last 40 years. These results highlight the importance of applying psychometric methods to the study of voice quality. Descriptive and correlational statistical techniques, including multidimensional scaling and factor analysis, may suggest that a given dimension is perceptually important. However, without confirmatory experimental studies, such associations between signal and percept remain merely suggestive. At a minimum, both theory and application will benefit if investigators verify that listeners are auditorily sensitive to the range of values of interest for a particular acoustic measure to ensure that such measures are in fact determinants of perceived quality.

ACKNOWLEDGMENTS

Norma Antoñanzas and Brian Gabelman wrote the software used in this research. Jason Mallory created the stimuli in experiment 2 and patiently tested many of the listeners in both experiments. Preliminary reports of the results were presented at the 144th and 145th meetings of the Acoustical Society of America. This research was supported by Grant DC01797 from the National Institute on Deafness and Other Communication Disorders.

¹Amplitude modulations did not produce major differences in the quality of most voices, because many so-called “amplitude tremors” are artifacts of frequency modulation (Sundberg, 1995; Kreiman *et al.*, 2003). However, these modulations were important for synthesizing several samples in the present study, especially those from speakers with spasmodic dysphonia.

²Software is available from the authors by request.

³Confidence ratings were not used in analyses of these data.

⁴Difference limens observed here may depend in part on the fact that listeners heard stimuli in free field, and could only hear a pair of voices once before responding. Although results reflect listeners’ discrimination ability under normal listening conditions, they do not necessarily reflect the precise limits of perceptual acuity as measured with headphones and multiple presentations prior to response.

Andrianopoulos, M. V., Darrow, K. N., and Chen, J. (2001). “Multimodal standardization of voice among four multicultural populations: Fundamental frequency and spectral characteristics,” *J. Voice* **15**, 194–219.

ANSI (1960). ANSI S1.1-1960, “Acoustical terminology” (American National Standards Institute, New York).

Bielamowicz, S., Kreiman, J., Gerratt, B. R., Dauer, M. S., and Berke, G. S. (1996). “A comparison of voice analysis systems for perturbation measurement,” *J. Speech Hear. Res.* **39**, 126–134.

Buder, E. H. (2000). “Acoustic analysis of voice quality: A tabulation of algorithms 1902–1990,” in *Voice Quality Measurement*, edited by R. D. Kent and M. J. Ball (Singular, San Diego), pp. 119–244.

Catford, J. C. (1977). *Fundamental Problems in Phonetics* (Indiana University Press, Bloomington, IN).

Coleman, R. F., and Wendahl, R. W. (1967). “Vocal roughness and stimulus duration,” *Speech Monographs* **34**, 85–92.

Deal, R., and Emanuel, F. W. (1978). “Some waveform and spectral features of vowel roughness,” *J. Speech Hear. Res.* **21**, 250–264.

de Krom, G. (1993). “A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals,” *J. Speech Hear. Res.* **36**, 254–266.

de Krom, G. (1995). “Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments,” *J. Speech Hear. Res.* **38**, 794–811.

Fant, G., Liljencrants, J., and Lin, Q. (1985). “A four-parameter model of glottal flow,” *STL-QPSR* **4**, 1–13.

Fukazawa, T., El-Assuooty, A., and Honjo, I. (1988). “A new index for evaluation of the turbulent noise in pathological voice,” *J. Acoust. Soc. Am.* **83**, 1189–1193.

Gabelman, B. (2003). “Analysis and synthesis of pathological vowels,” unpublished doctoral dissertation, Department of Electrical Engineering, University of California at Los Angeles.

Gerratt, B. R., and Kreiman, J. (2001). “Measuring voice quality with speech synthesis,” *J. Acoust. Soc. Am.* **110**, 2560–2566.

Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics* (Krieger, Huntington, NY).

Hanson, D. G., Jiang, J. J., Chen, J., and Pauloski, B. R. (1997). “Acoustic measurement of change in voice quality with treatment for chronic posterior laryngitis,” *Ann. Otol. Rhinol. Laryngol.* **106**, 279–285.

Hartl, D. A., Hans, S., Vaissiere, J., and Brasnu, D. A. (2003). “Objective acoustic and aerodynamic measures of breathiness in paralytic dysphonia,” *Eur. Arch. Otorhinolaryngol.* **260**, 175–182.

Heiberger, V. L., and Horii, Y. (1982). “Jitter and shimmer in sustained phonation,” in *Speech and Language: Advances in Basic Research and Practice*, edited by N. J. Lass (Academic, New York), Vol. 7, pp. 299–332.

Hillenbrand, J. (1987). “A methodological study of perturbation and additive noise in synthetically generated voice signals,” *J. Speech Hear. Res.* **30**, 448–461.

Hillenbrand, J. (1988). “Perception of aperiodicities in synthetically generated voices,” *J. Acoust. Soc. Am.* **83**, 2361–2371.

Javkin, H., Antonanzas-Barroso, N., and Maddieson, I. (1987). “Digital inverse filtering for linguistic research,” *J. Speech Hear. Res.* **30**, 122–129.

Jensen, P. J. (1965). “Adequacy of terminology for clinical judgment of voice quality deviation,” *Eye Ear Nose Throat Mon.* **44**, 77–82.

Jones, T. M., Trabold, M., Plante, F., Cheetham, B. M., and Earis, J. E. (2001). “Objective assessment of hoarseness by measuring jitter,” *Clin. Otolaryngol.* **26**, 29–32.

Jotz, G. P., Cervantes, O., Abrahao, M., Settanni, F. A., and de Angelis, E. C. (2002). “Noise-to-harmonics ratio as an acoustic measure of voice disorders in boys,” *J. Voice* **16**, 28–31.

Karnell, M. P., Hall, K. D., and Landahl, K. L. (1995). “Comparison of fundamental frequency and perturbation measurements among three analysis systems,” *J. Voice* **9**, 383–393.

Kreiman, J., and Gerratt, B. R. (1996). “The perceptual structure of pathologic voice quality,” *J. Acoust. Soc. Am.* **100**, 1787–1795.

Kreiman, J., and Gerratt, B. R. (2000). “Measuring vocal quality,” in *Voice Quality Measurement*, edited by R. D. Kent and M. J. Ball (Singular, San Diego), pp. 73–102.

Kreiman, J., Gabelman, B., and Gerratt, B. R. (2003). “Perception of vocal tremor,” *J. Speech Lang. Hear. Res.* **46**, 203–214.

Kreiman, J., Gerratt, B. R., and Berke, G. S. (1994). “The multidimensional nature of pathologic vocal quality,” *J. Acoust. Soc. Am.* **96**, 1291–1302.

Kreiman, J., Gerratt, B. R., and Precoda, K. (1990). “Listener experience and perception of voice quality,” *J. Speech Hear. Res.* **33**, 103–115.

Kreiman, J., Vanlancker-Sidtis, D., and Gerratt, B. R. (2005). “Perception of voice quality,” to appear in *Handbook of Speech Perception*, edited by D. B. Pisoni and R. E. Remez (Blackwell, Oxford), pp. 338–362.

Linville, S. E. (2000). “The aging voice,” in *Voice Quality Measurement*, edited by R. D. Kent and M. J. Ball (Singular, San Diego), pp. 359–376.

MacMillan, N. A., and Creelman, C. D. (1991). *Detection Theory: A User’s Guide* (Cambridge University Press, Cambridge).

Marks, L. E., and Algom, D. (1998). “Psychological scaling,” in *Measurement, Judgment, and Decision Making*, edited by M. H. Birnbaum (San Diego, Academic), pp. 81–178.

Martin, D., Fitch, J., and Wolfe, V. (1995). “Pathologic voice type and the acoustic prediction of severity,” *J. Speech Hear. Res.* **38**, 765–771.

Michaelis, D., Gramss, T., and Strube, H. W. (1997). “Glottal-to-noise excitation ratio—A new measure for describing pathological voices,” *Acustica* **83**, 700–706.

- Munoz, J., Mendoza, E., Fresneda, M. D., Carballo, G., and Lopez, P. (2003). "Acoustic and perceptual indicators of normal and pathological voice," *Folia Phoniatr Logop* **55**, 102–114.
- Murphy, P. J. (1999). "Perturbation-free measurement of the harmonics-to-noise ratio in voice signals using pitch synchronous harmonic analysis," *J. Acoust. Soc. Am.* **105**, 2866–2881.
- Niedzielska, G. (2001). "Acoustic analysis in the diagnosis of voice disorders in children," *Int. J. Pediatr. Otorhinolaryngol.* **57**, 189–193.
- Prosek, R. A., Montgomery, A. A., Walden, B. E., and Hawkins, D. B. (1987). "An evaluation of residue features as correlates of voice disorders," *J. Commun. Disord.* **20**, 105–117.
- Su, C. Y., Lui, C. C., Lin, H. C., Chiu, J. F., and Cheng, C. A. (2002). "A new paramedian approach to arytenoid adduction and strap muscle transposition for vocal fold medialization," *Laryngoscope* **112**, 342–350.
- Sundberg, J. (1995). "Acoustic and psychoacoustic aspects of vocal vibrato," in *Vibrato*, edited by P. H. Dejonckere, M. Hirano, and J. Sundberg (Singular, San Diego, CA), pp. 35–62.
- Titze, I. R. (1994). *Principles of Voice Production* (Prentice Hall, Englewood Cliffs, NJ).
- Titze, I. R. (1995). *Workshop on Acoustic Voice Analysis Summary Statement* (National Center for Voice and Speech, Denver).
- Titze, I. R., Horii, Y., and Scherer, R. C. (1987). "Some technical considerations in voice perturbation measurements," *J. Speech Hear. Res.* **30**, 252–260.
- Wendahl, R. W. (1963). "Laryngeal analog synthesis of harsh voice quality," *Folia Phoniatr.* **15**, 241.
- Wendahl, R. W. (1966a). "Some parameters of auditory roughness," *Folia Phoniatr.* **18**, 26–32.
- Wendahl, R. W. (1966b). "Laryngeal analog synthesis of jitter and shimmer auditory parameters of harshness," *Folia Phoniatr.* **18**, 98–108.

Consonant recognition and the articulation index

Jont B. Allen

ECE Department and the Beckman Institute, University of Illinois, Urbana, Illinois 61801

(Received 1 August 2004; revised 14 December 2004; accepted 15 December 2004)

The purpose of this paper is to provide insight into how speech is processed by the auditory system, by quantifying the nature of nonsense speech sound confusions. (1) The Miller and Nicely [J. Acoust. Soc. Am. **27**(2), 338–352 (1955)] confusion matrix (CM) data are analyzed by plotting the CM elements $S_{i,j}$ (SNR) as a function of the signal-to-noise ratio (SNR). This allows for the robust clustering of perceptual feature (event) groups, not robustly defined by a single CM table, where clusters depend on the sound order. (2) The SNR is then re-expressed as an articulation index (AI), and used as the independent variable. The normalized log scores $\log(1-S_{i,i}(\text{AI}))$ and $\log(S_{i,j}(\text{AI}))$, $j \neq i$, then become linear functions of AI, on log-error versus AI plots. This linear dependence may be interpreted as an extension of the band-independence model of Fletcher. (3) The model formula for the average score for the finite-alphabet case $P_c(\text{AI}, \mathcal{H}) = \sum_{i=1}^N S_{i,i} / N$ is then modified to include the effect of entropy \mathcal{H} . Due to the grouping of sounds with increased SNR (and AI), the sound-group entropy \mathcal{H}_g plays a key role in this performance measure. (4) A parametric model for the confusions $S_{i,j}(\text{AI}, \mathcal{H}_g)$ is then described, which characterizes the confusions between competing sounds within a group. © 2005 Acoustical Society of America.
[DOI: 10.1121/1.1856231]

PACS numbers: 43.71.Gv, 43.71.An, 43.72.Ne [DOS]

Pages: 2212–2223

I. INTRODUCTION

The articulation index (AI) and the confusion matrix (CM), denoted $\mathcal{C}_{i,j}$, are two important measures frequently used to characterize human speech recognition. This paper explores the merging of these two measures by expressing the CM as a function of the AI. The natural grouping of sounds, observed by Miller and Nicely in their classic 1955 experiment, is then explained in terms of a change in each group's entropy \mathcal{H}_g , which is also a function of the AI. A parametric model of the symmetric form of the CM confusions $S_{i,j}(\text{AI}, \mathcal{H})$ is then developed. These parametric models lead to a new formulation of the average score $P_c(\text{AI}, \mathcal{H})$, which accounts for chance. This model provides an accurate fit to the raw data, and provides new insight into Fletcher's *band independence* formulation for the average articulation score (Allen, 1994).

A fundamental building block in the theory of speech communication is the articulation index (AI) model of speech sound recognition (French and Steinberg, 1947; Fletcher and Galt, 1950; Allen, 1994). The AI is a speech audibility measure, averaged across many cochlear frequency bands, of a specific speech-to-noise ratio measure, that takes into account the effects of masking and cochlear filtering (Allen, 1994). The masking can be due either to external noise, when noise is added to the speech at the source, or internal noise, present in the cochlea and auditory nerve. The AI measure is weighted to account for both cochlear critical bands and those frequency regions containing important speech information. Fletcher developed the AI to predict the average nonsense-phone articulation score $P_c(\text{SNR})$, as a function of the speech-to-noise ratio (SNR) (specified over 20 frequency bands), as a way of avoiding expensive and time-consuming listening tests on speech communication equipment. The AI also offers important in-

sights into human speech perception. In this paper the AI is used to provide a new and insightful view of the confusion data of Miller and Nicely (1955).

Many procedures have been developed which claim to predict the phone performance score for noisy and/or filtered channels, several of which are called the “articulation index.” Following the first proposal in 1921 by Fletcher, there is the Bell Labs procedure of FS47 (key abbreviations are provided in Table I), followed by the more extensive version by Fletcher and Galt (1950). Then, Kryter published his simplified method in 1962 (Kryter, 1962a, 1962b), soon followed by the ANSI version. More recently, the STI was proposed to extend the AI procedure when room reverberation is present (Houtgast and Steeneken, 1973; Steeneken and Houtgast, 1980). This was then followed by the new ANSI procedure SII (S3.5-1997, 1997).

The focus on the articulation index provided here is not on predicting the performance of speech communications systems, but rather in understanding and modeling the perception and recognition of human speech sounds. The focus

TABLE I. Table of abbreviations.

Abbreviation	Definition
CV, CVC	Consonant+vowel sounds
intelligibility	Recognition of phonemes
articulation	Recognition of nonsense phones
CM	Confusion matrix
AM	Articulation matrix (CM of nonsense speech)
AI	Articulation index
SNR	Speech-to-noise ratio
rms	Root-mean-squared
events	Perceptual features
MN55	Miller and Nicely (1955)
FS47	French and Steinberg (1947)

here is not on procedures for predicting phone intelligibility, but in gaining leverage from the well-validated AI predictions, to provide insight into nonsense-phone identifications.

At least four fundamental questions are addressed in this presentation.

- (1) What is the relation between the AI and the CM?
- (2) How can one more accurately determine the groups that are present in Miller and Nicely's CM data at certain SNRs?
- (3) What are the limits of Fletcher's *band independence* assumption?
- (4) Can the AI be used to predict the phone score for closed sets (i.e., what are the limits of the AI procedure)?

Subsequent to the Bell Labs articulation studies, and following up work done during WWII, Miller took up the study of speech articulations in much greater detail. As detailed in Miller's (1951) book *Language and Communication*, information and communication theory are the basis for understanding the speech code. One of the basic tools of information theory is the *channel*, the mathematical characterization of a communication link (i.e., a noisy pair of wires, with codecs attached), in terms of discrete input and output symbols from some *alphabet*. One method for characterizing the human speech communication channel is the nonsense-phone confusion matrix (CM), which characterizes the probabilities of nonsense speech sound (the symbols) transmission errors. It is expected that an error analysis of this *articulation matrix* (AM), as a function of the SNR, can give important insight into the speech code.

A second key concept from information theory is that of entropy \mathcal{H} , which is a measure of the compactness of a probability distribution, which in the case of speech represents the distribution of the sound confusions. In his classic 1951 study, Miller *et al.* (1951) (MHL51) showed the effects of symbol alphabet size, and thus the entropy, on word recognition. Four years later, in a second classic study, Miller and Nicely (MN55) showed that as the wideband SNR is raised from -18 to $+12$ dB, *the sounds form perceptual groups*. The formation of a group, as a function of the SNR, also results in a change (reduction) in \mathcal{H} .

In analyzing their confusion matrices, MN55 quantified the grouping effect using a mutual-information (MI) analysis, on assumed groups. An natural advantage of the MI analysis method is its insensitivity to bias, as defined by the skew-symmetric form of the CM. This is at the same time a weakness of the MI, since it may be beneficial to remove the effects of subject bias prior to modeling the confusions. A major disadvantage of mutual information, as used in MN55, is that it gives no insight into the formation of the groups being analyzed.

This paper explores the limits and applicability of Fletcher's band-independence model, applied to the 1955 closed-set consonant articulation data of Miller and Nicely (MN55). Such data do not meet the usual assumptions of the AI audibility measure, of a large, high-entropy open-set corpus. By plotting the confusions as a function of SNR, it is possible to identify the groups in a systematic, logical way, without assuming any predetermined sound ordering. This

leads to an accurate method of identifying the natural sound groups, and allows one to display the complex body of CM data at all SNRs, in a single figure. Once the groups have been identified, one is then free to further explore the relationship of the AM to the AI and \mathcal{H} .

The definitions of mathematical symbols has been summarized in Table II.

II. REPRESENTATIONS OF THE CONFUSION MATRIX (CM)

Figure 1 shows a typical MN55 consonant–vowel (CV) *confusion matrix* or *count matrix* for wideband speech (0.2–6.5 kHz), at a *speech-to-noise ratio* (SNR) of -6 dB (Miller and Nicely, 1955, Table III). The 16 consonants were presented along with the vowel /a/ as in father (i.e., the first three sounds were [pa/,ta/,ka/]). After hearing one of the 16 CV sounds as labeled by the first column, the consonant that was reported is given as labeled along the top row. This array of numbers form the basic CM, denoted $\mathcal{C}_{s,h}$, where integer indices s and h (i.e., “spoken” and “heard”) each run between 1 and 16. For example, /pa/ was spoken 230 times (the sum of the counts in the first row), and was reported heard 80 times ($\mathcal{C}_{1,1}$), while /ta/ was reported 43 times ($\mathcal{C}_{1,2}$). For Table III the mean row count was 250, with a standard deviation of 21 counts.

When the sounds are ordered as shown in Fig. 1, they form groups, identified in terms of hierarchical clusters of *articulatory features*. For example, the first group of sounds 1–7 correspond to unvoiced, group 8–14 are voiced, and 15, 16 are nasal (and also voiced).

At an SNR of -6 dB, the intraconfusions (within a group) are much greater than the interconfusions (between groups). For example, members of the group 1–7 (the unvoiced sounds) are much more likely to be confused among themselves, than between the voiced sounds (8–14), or the nasal sounds (15,16). The nasal are confused with each other, but rarely with any of the other sounds 1–14.

TABLE II. Table of mathematical symbols.

Symbol	Definition	Equation
\mathcal{I}	Intelligibility (meaningful sound recognition)	
\mathcal{C}	Confusion matrix	(2)
\mathcal{A}	Articulation matrix	(1)
S	Symmetric form of \mathcal{A}	(3)
A	Skew-symmetric form of \mathcal{A}	(4), (5)
P_c	Probability correct	(17), (16), (18)
$P_c^{(i)}$	Same as $\mathcal{A}_{i,i}$ and $S_{i,i}$	
\mathcal{H}	Entropy	
AI	Articulation index (AI)	(8)
AI_k	Specific AI in band k	(11)
e	Total error	(12), (13), (20)
e_i	Total error for sound i : $e_i \equiv 1 - P_c^{(i)} = 1 - \mathcal{A}_{i,i}$	(7)
e_{\min}	Minimum error	(15)
e_{chance}	Chance error	(19)
ϵ_k	k th band error	(14)
$e_{\min}^{(i)}$	Minimum error for sound i : $e_i _{AI=1}$	
S_{15}	Shorthand for $S_{15,15}$	
snr_k	Speech-to-noise rms ratio in band k	(9)

TABLE III. Confusion matrix for $S/N = -6$ db and frequency response of 200–6500 cps.

	<i>p</i>	<i>t</i>	<i>k</i>	<i>f</i>	<i>θ</i>	<i>s</i>	<i>ʃ</i>	<i>b</i>	<i>d</i>	<i>g</i>	<i>v</i>	<i>ð</i>	<i>z</i>	<i>ʒ</i>	<i>m</i>	<i>n</i>	
STIMULUS	<i>p</i>	80	43	64	17	14	6	2	1	1	1	1	2		2		
	<i>t</i>	71	84	55	5	9	3	8	1			1	2		2	3	
	<i>k</i>	66	76	107	12	8	9	4				1			1		
	<i>f</i>	18	12	9	175	48	11	1	7	2	1	2	2				
	<i>θ</i>	19	17	16	104	64	32	7	5	4	5	6	4	5			
	<i>s</i>	8	5	4	23	39	107	45	4	2	3	1	1	3	2		1
	<i>ʃ</i>	1	6	3	4	6	29	195		3							1
	<i>b</i>				5	4	4		136	10	9	47	16	6	1	5	4
	<i>d</i>							8	5	80	45	11	20	20	26	1	
	<i>g</i>					2			3	63	66	3	19	37	56		3
	<i>v</i>				2		2		48	5	5	145	45	12		4	
	<i>ð</i>					6			31	6	17	86	58	21	5	6	4
	<i>z</i>					1	1	1	7	20	27	16	28	94	44		1
	<i>ʒ</i>								1	26	18	3	8	45	129		2
	<i>m</i>	1							4			4	1	3		177	46
	<i>n</i>					4			1	5	2		7	1	6	47	163

FIG. 1. Typical Miller–Nicely frequency of confusions, or count matrix C , from Table III at -6 -dB SNR. Each entry in the matrix $C_{s,h}$ is the subject response count. The rows correspond to the *spoken* CVs, each row representing a different consonant, from $s = 1, \dots, 16$. The columns correspond to the *heard* CVs, each column representing a different consonant, from $h = 1, \dots, 16$. The common vowel /a/, as in “father,” was used throughout. When the 16 consonants are ordered as shown, the count matrix shows a “block-symmetric” partitioning in the consonant confusions. In this matrix there are three main blocks delineated by the dashed lines, corresponding to unvoiced, voiced, and nasal. Within the voiced and unvoiced subgroups, there are two additional symmetric blocks, corresponding to affrication and duration, also delineated with dashed lines.

The MN55 articulatory feature classification scheme is far from perfect. For example, the nasals are voiced in the same sense as those labeled voiced; however, they clearly form a unique cluster. Thus, there is no unique simple articulatory label for sounds 8–14. Groups systematically depend on the SNR, and groups remain unidentified by this scheme. Using the example of Table III (Fig. 1), [ba/,va/,ða/] form a group that is distinct from the nonfricative voiced subgroup. An improved order for sounds 8–14 would be [ba/,va/,ða/], [za/,za/,da/,ga/]. Of course, this example fundamentally breaks the MN55 articulatory feature classification scheme. In fact, the feature space cannot strictly be articulatory feature based.

The MN55 data have been the inspiration for a large number of studies. The sound grouping has been studied using multidimensional scaling, which has generally failed in providing a robust method for finding perceptually relevant groups of sounds, as discussed by Wang and Bilger (1973). Thus, the grouping problem has remained unsolved.

The data in the CM represent a psychological subject response, and therefore need to be represented in terms of *psychological variables* rather than physical (production) measures, as labeled by articulatory features. This could have been the role of *distinctive features*, had they been so defined. Unfortunately, there seems to be some confusion in the literature as to the precise definition of a distinctive feature. For example, are distinctive features production or perception quantities?

To avoid this confusion, I shall use the term *event* when referring to *perceptual features*. Since Miller and Nicely’s confusion data are based on perception, they must be de-

scribed by events. The precise nature of these events may be explored by studying the 15 plots $S_{i,j}$, $i \neq j$, as shown in the lower-left panel of Fig. 2 for the case of $i = 2$.

A. The transformation from CM to AM

The term *articulation* is defined as the probability correct P_c of identifying nonsense-phone speech sounds (consonants and vowels), while *intelligibility* \mathcal{I} is the probability of identifying meaningful speech sounds, such as words and sentences (Fletcher and Galt, 1950).

When normalized as a probability, the consonant confusion matrix is transformed to an *articulation matrix* (AM), denoted \mathcal{A} (script A, Table II), with elements

$$A_{s,h} \equiv \frac{C_{s,h}}{\sum_h C_{s,h}}. \quad (1)$$

This normalization, to an equal probability for each row, is justified because of the small standard deviation of the row sums (i.e., 250 ± 21).

The AM is the empirical conditional probability $P_c(h|s)$ of reporting sound h after speaking sound s , namely

$$A_{s,h} \equiv P_c(h|s), \quad (2)$$

for integer labels s, h (i.e., spoken, heard). In another sense, $A_{s,h}$ for $s \neq h$ is an error probability, since it is the probability of reporting the wrong sounds h after hearing spoken sound $s \neq h$.

Figure 2 shows the probability of responding that the sound $h = 1, \dots, 16$ was reported, following speaking /ta/ ($s = 2$), as a function of the wideband SNR. The upper-left

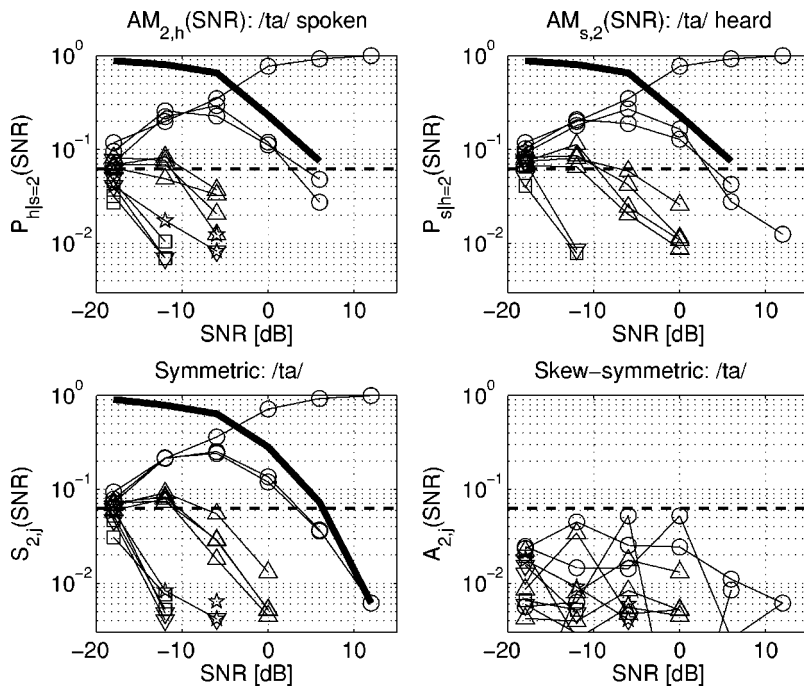


FIG. 2. This figure shows Miller and Nicely's 1955 wideband row-normalized confusion matrix data $\mathcal{A}_{s,h}(\text{SNR})$ [Eq. (1)] for the sound /ta/ (sound 2) from MN55 Tables I–IV, as a function of the speech-to-noise ratio. The upper-left panel is a plot of the second row of the articulation matrix $[\mathcal{A}_{2,h}(\text{SNR}), h=1,\dots,16]$, corresponding to /ta/ spoken], while the upper-right panel is a plot of the second column $[\mathcal{A}_{s,2}(\text{SNR})]$, corresponding to /ta/ heard]. The matrix is not perfectly symmetric ($\mathcal{A} \neq \mathcal{A}^t$), which explains the small differences between these two plots. The lower-left panel is the symmetric form of the articulation matrix given by Eq. (3), which is the average of \mathcal{A} and its transpose \mathcal{A}^t . The lower-right panel is the skew-symmetric form A [Eq. (4)]. The horizontal dashed line in each figure shows chance performance (i.e., $1/16$).

panel shows the probability $\mathcal{A}_{2,h}(\text{SNR})$ of each heard sound ($h=1,\dots,16$), given /ta/ was spoken. The upper-right panel shows the probability $\mathcal{A}_{s,2}$ of each sound spoken ($s=1,\dots,16$), given that /ta/ was heard. The curve that rises to 1 is the probability of correctly reporting /ta/ $\mathcal{A}_{2,2}(\text{SNR})$, given that it was spoken (left), or spoken given that it was heard (right). The solid-thick curve is the total probability of error $e_2(\text{SNR}) \equiv 1 - \mathcal{A}_{2,2}(\text{SNR})$ of not reporting /ta/, given that it was spoken (left) or heard (right).

Symmetric and skew-symmetric decomposition

The lower-left panel of Fig. 2 is a plot of the second row $S_{2,j}$ of the symmetric form of the AM, defined as

$$S \equiv \frac{1}{2}(\mathcal{A} + \mathcal{A}^t), \quad (3)$$

where \mathcal{A}^t is the transpose of \mathcal{A} , while the lower-right panel is the second row of $A_{i,j}$ of the skew-symmetric form of the matrix, defined as

$$A \equiv \frac{1}{2}(\mathcal{A} - \mathcal{A}^t). \quad (4)$$

It appears that the sampling error (statistical uncertainty) in the measurements, due to the sample size, is about 0.5% (0.005), which is where the measurements become scattered. This variability is determined by many factors, including the number of trials per sound, the smoothing provided by the symmetric transformation, the consistency of the talker, and the mental concentration and number of the observers (four in this case).

From the lower-right panel, it is clear that the AM is close to symmetric, since the skew-symmetric terms are small. A few terms of $A_{2,h}(\text{SNR})$ are as large as 5%, but most are less than 1%. Since the MN55 data are close to symmetric, it is reasonable to force the symmetry, and then to study S and A separately, which is the approach taken here. Note that S is slightly smoother than \mathcal{A} , since each element $\mathcal{A}_{s,h}$ is the average of two similar terms, $\mathcal{A}_{h,s}$ and

$\mathcal{A}_{s,h}$. Using the symmetric form simplifies the analysis of the matrix and gives us access to the skew-symmetric form.

Based on an analysis by Goldstein (1980), the interpretation of the skew-symmetric form is quite different from that of the symmetric form. The most likely explanation of the skew-symmetric matrix is that the subjects have a bias for one sound over another, and are therefore more likely to report the consonant for which they have the bias (Goldstein, 1980).

The largest skew-symmetric sounds in row 2 are /fa/, /θa/, /va/, and /ða/, which have errors approaching 5%, but are always less than chance (1/16). It seems significant that the skew-symmetric form always lies slightly below chance (Fig. 2, lower-right panel). For the rest of the sounds, the error patterns are similar in their nature to those of /ta/, with the largest errors of about 10% in a few places, but with most of the errors being a few percent or less.

There is an interaction between the row normalization [Eq. (1)], and the symmetry transformation Eq. (3), which requires that the row normalization and symmetric computations be iterated. This iteration always converges to the same result, and is always stable for all of the MN55 tables. An entry of "1" in \mathcal{C} represents a single vote for the same utterance, from four listeners who heard that utterance. All 1's were deleted from the matrix before computing S . Once matrix S has been determined, A is computed from

$$A = \mathcal{A} - S. \quad (5)$$

Plotting the symmetric data $S(\text{SNR})$ as a function of SNR, as shown in Fig. 2, provides a concise yet comprehensive summary of the entire set of measurements, and shows the hierarchical grouping, without a need to order the sounds. In the next section it is shown that if $S_{i,j}$ is described as a function of the AI, rather than the SNR, the same data may be quantitatively modeled, and the important effects of chance may be accounted for.

B. Grouping the sounds

Sound clustering in the CM was used by MN55 as the basis for arguing that the sounds break down into distinct groups, which MN55 identified as five discrete *articulatory features*, which they called *voicing*, *nasality*, *affrication*, *duration*, and *place*.

Each symbol in Fig. 2 labels a different articulatory feature. Sounds 1–3 (/pa/,/ta/,/ka/) are shown as circles, 4–7 (/fa/,/θa/,/sa/,/ʃa/) triangles, 8–10 (/ba/,/da/,/ga/) squares, 11–14 (/va/,/ða/,/za/,/ʒa/) upside-down triangles, while the nasal sounds 14 and 15 (/ma/,/na/) are labeled by 5-pointed stars.

The hierarchical clusters are seen as groups that peel away as the SNR increases. The symmetric /ta/ data shown in the lower-left panel of Fig. 2 are a great example: First, all the voiced sounds dramatically drop, starting from chance, as the SNR is raised. Next, the unvoiced-fricatives /fa/, /θa/, /sa/, /ʃa/ (triangles) peel off, after very slightly rising above chance at –12-dB SNR. Finally, the two main competitors to /ta/ (/pa/ and /ka/) peak around –6-dB SNR, and then fall dramatically, as /ta/ is clearly identified at 0-dB SNR and above. In the lower-left panel /pa/, /ta/, and /ka/ (○) are statistically indistinguishable below –6 dB, and approach chance identification of 1/16 at –18 dB. Above about –6 dB, /ta/ separates and the identification approaches 1, while the confusions with the other two sounds (/pa/ and /ka/) reach a maximum of about a 25% score, and then drop monotonically, as the SNR increases.

The MN55 sounds 4–7 (/fa/, /θa/, /sa/, and /ʃa/), like sounds 1–3 (/pa/,/ta/,/ka/), also form a group, as may be seen in the lower-left panel, labeled by Δ. This group also starts from chance identification (6.25%), rises slightly to a score of about 7% at –12 dB, and then monotonically drops at a slightly greater rate than sounds 1 and 3 (symbols ○).

The third group is the remaining sounds 8–16, labeled by the remaining symbols, which show no rise in performance; rather, they steeply drop from the chance level.

At the lowest SNR of –18 dB, the elements in the symmetric form of the AM approach chance performance, which for MN55 is 1/16, corresponding to closed-set guessing. Extrapolating the data of Fig. 2, chance performance corresponds to about –21-dB SNR.

Based on the clustering seen in the AM (e.g., MN55 Tables II and III), it was concluded by MN55 that the three sounds /ta/, /pa/, and /ka/ might be thought of as one group. These three sounds form the unvoiced, non-nasal, nonaffricate, low-duration group, having three different values of place. The details of these groupings depend on the SNR. A detailed analysis of these clusters show that the MN55 *articulatory features* (production feature set) do not always correspond to the *events* (perceptual feature set).

In fact, it would be surprising if it turned out any other way, given that production and perception are fundamentally different things. The details of a scheme that will allow us to make such an analysis of the optimal perceptual feature set, form the remainder of this paper.

1. Formula for the total error

The solid-thick curve in the top two, and bottom-left panels of Fig. 2, are graphs of the total error for /ta/

$$e_2(\text{SNR}) \equiv 1 - S_{2,2}(\text{SNR}). \quad (6)$$

Because each row of $S_{i,j}$ has been normalized so that it sums to 1, the total error for the i th sound is also the row sum of the 15 off-diagonal ($j \neq i$) elements, namely

$$e_i(\text{SNR}) = \sum_{j \neq i} S_{i,j}(\text{SNR}). \quad (7)$$

Since each error term is non-negative, e_i must bound each individual confusion $S_{i,j}$. For the data of Fig. 2, lower-left, the other two circle curves (/pa/ and /ka/), which compete with /ta/, and thereby form a 3-group, are nearly identical. All other error terms are much smaller. Thus, the solid-thick curve, $e_2(\text{SNR})$, is approximately twice the size of the curves for /pa/ and /ka/. All the off-diagonal terms go to zero at +12-dB SNR so for that one point $e_2 = S_{2,j}$, a fluke of the small-number statistics.

Equation (7) says that the total error for the i th sound is linearly decomposed by the off-diagonal errors of the AM. This is a natural decomposition of the total error into its confusions that can help us understand the AI predictions in much greater detail.

For example: *Why does the probability of identification of sounds 1–3 and 4–7 increase even when these sounds are not spoken?* The initial rise for the two sound groups follows from the increase in chance performance due to the decreased entropy, which follows from the reduced size of the group. This conclusion follows naturally from Eq. (7). As the SNR increases, the size of the group exponentially decreases.

As the number of alternatives in a closed-set task decreases, the probability of guessing increases. Given 2 alternatives, chance is 1/2; given 16, chance is 1/16. Thus, grouping and the rise due to the confusion within the group are intimately tied together. In the same manner, as the SNR rises from –18 to –12, the MN55 sounds 4–16 are perceptually ruled out, increasing chance performance for sounds 1–3 from 1/16 to 1/3.

2. The nasals

In Fig. 3 $S_{i,j}(\text{SNR})$ for $i = 15, 16$, corresponding to /ma/ and /na/, are presented. The two nasal sounds are clearly separated from all the other sounds, even at –18-dB SNR. As the SNR increases, the scores rise to $\approx 25\%$, peaking at or near –12-dB SNR, following with the identification rising and the confusion dramatically falling for SNRs at and above –6 dB.

Sounds 1–14 are solidly rejected, even at –18 dB. These scores exponentially drop as the SNR is increased. There is a slight (visual) hint of a rise of a few sounds for the case of /ma/, in some of the rejected sounds in the left panel, and some corresponding grouping, but the effect is small and it would be difficult to tease out. The rejected sounds in the right panel do not show any obvious grouping effect.

The subjects can clearly distinguish the two nasal sounds (sounds 15,16) from all the others (sounds 1–14),

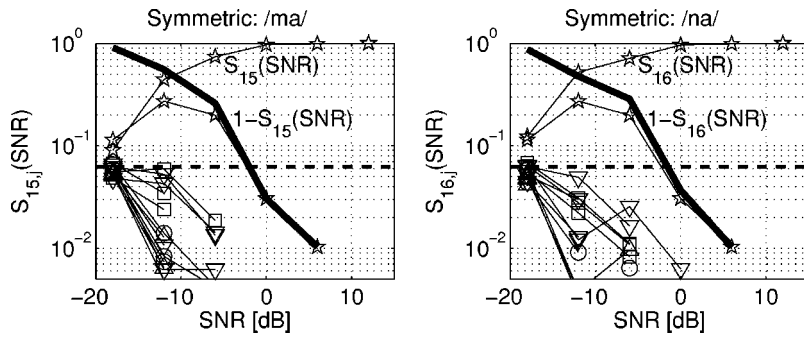


FIG. 3. Plots of the symmetric AM corresponding to the nasals /ma/ and /na/. The curve that rises to 1 is $S_{i,i}(\text{SNR})$ for $i=15$ (left) and $i=16$ (right). The solid thick curve in each panel is e_i [Eq. (7)]. The other curves represent confusions $S_{i,j}(\text{SNR})$ for the remaining sounds $j=1, \dots, 14$.

even at the lowest SNR of -18 dB; however, they cannot distinguish between them until the SNR is greater than -12 dB. The subjects know the sound they hear is nasal, but the question is, which one? This identification of event-nasal leads to a significant increase in chance performance for SNRs between -18 and -6 dB, from $1/16$ to $1/2$.

One may also see this effect in the raw count data at -18 dB, where the confusions are approaching equal chance levels. For example, in MN55 Table I, the raw counts are [25,28;33,32]. At -12 dB, /ma/ and /na/ are significantly confused with each other, but rarely with the other sounds. For example, from MN55 Table II, /ma/ is heard 20 times when /ba/ is spoken [$S_{15,8}(-12) = 6.72\%$ of the time], while /ba/ is heard 11 times when /ma/ is spoken (5.83% of the time).

III. TRANSFORMATION FROM THE WIDEBAND SNR TO THE AI

Miller and Nicely used the wideband SNR, in dB, as their measure of audibility. However, as discussed in the Introduction, there are reasons to believe that the AI(SNR) is a better audibility measure. We shall now demonstrate this for the MN55 data. Our approach is to transform MN55's wideband SNR into an AI, and then to plot the resulting $S_{i,j}(\text{AI})$.

To compute the AI for MN55 one needs to know the *specific* SNR, over articulation bands, denoted snr_k . This requires knowledge of the average speech spectra for five female talkers, and the noise spectra. The spectrum for five female talkers is shown in Fig. 4, while the noise spectra was independent of frequency (i.e., white). The procedure for computing AI(SNR) is described next.

A. Computing the specific AI

The AI is defined by FS47 [their Eq. (8)] as

$$\text{AI} = \frac{1}{K} \sum_k^K \text{AI}_k, \quad (8)$$

namely as a 20-band average over the *specific* AI, denoted AI_k . The specific AI is defined in terms of the speech-to-noise ratio

$$\text{snr}_k \equiv \sigma_{s,k} / \sigma_{n,k}, \quad (9)$$

where the speech power is $\sigma_{s,k}^2$ [Watts/critical band] and the masking noise power is $\sigma_{n,k}^2$ [Watts/critical band], in the k th articulation band. When calculating $\sigma_{s,k}$, the average is over 1/8-s intervals. snr_k is the same as FS47's band sensation

level E. The k th articulation band power-snr *speech detection threshold* may be modeled as

$$\frac{I + \Delta I}{I} \equiv \frac{\sigma_{n,k}^2 + c^2 \sigma_{s,k}^2}{\sigma_{n,k}^2} = 1 + c^2 \text{snr}_k^2, \quad (10)$$

where a frequency-independent *speech detection constant* c is determined empirically from data on the detection of speech in noise (Fletcher and Munson, 1937; French and Steinberg, 1947). The role of c is to convert the speech rms to the speech peaks, which are typically 12 dB above the rms speech level. When snr_k specifies the speech peaks, $c=2$.

Converting to decibels, and scaling by 30, defines the *specific* AI

$$\text{AI}_k = \min\left(\frac{1}{3} \log_{10}(1 + c^2 \text{snr}_k^2), 1\right). \quad (11)$$

Relationship Eq. (11) follows from the detailed discussions of FS47 and Fletcher and Galt (1950), followed by the subsequent analysis by Allen (1994). [See especially (Fletcher, 1995, Eq. (10-3), page 167).]

Between 0 and 30 dB, AI_k is proportional to $\log(\text{snr}_k)$ because the percent of the time the speech is above a certain level is proportional to the dB SL level (*re*: threshold sensation level) (French and Steinberg, 1947; Allen, 1994). The factor of $1/3$ comes from the dynamic range of speech in a given articulation band (French and Steinberg, 1947, Fig. 4,

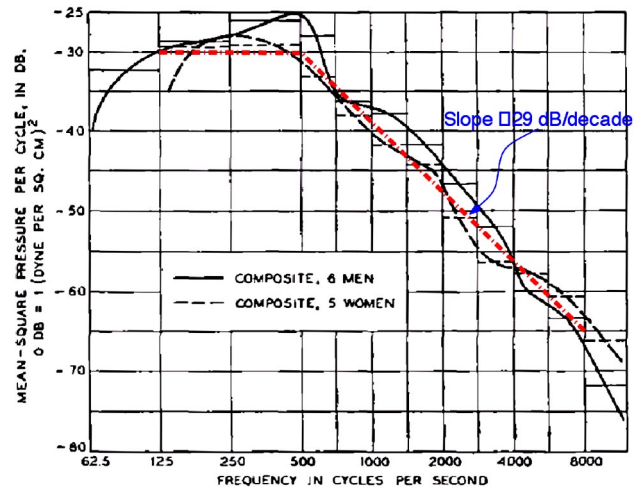


FIG. 4. This figure from Dunn and White, 1940 (their Fig. 10) shows the average power spectrum for six men and five women. The dashed curve, which approximates the power spectrum for the five women, has a slope of 0 from 125 to 500 Hz, and a slope of -29 -dB/decade between 0.5 and 8 kHz.

page 95). As discussed extensively by FS47 (i.e., their Table 12), an empirical *threshold adjustment* must be made, labeled c in Eq. (10). The value of c is chosen such that the speech is just detectable when $\text{snr}_k = 1$, in each cochlear critical band, corresponding to specific AIs of zero (i.e., $\text{AI}_k = 0$). Equation (11) over-predicts the data of FS47 Table V by 3.2% ($c \equiv 1/2$, $\text{snr}_k = E$). A more precise estimation of c would require repeating Fletcher's critical ratio experiment using narrow bands of speech, with a white-noise masker and measuring snr_k at the detection threshold. The $\min(x, 1)$ part of the definition limits the AI on the high end, since for an SNR above 30 dB, the noise has a negligible effect on the articulation (and intelligibility).

The band independence model of the total error

The average sound articulation error e (SNR), in terms of the average sound articulation $P_c(\text{SNR})$, is

$$e(\text{SNR}) = 1 - P_c(\text{SNR}). \quad (12)$$

In 1921 Fletcher showed that the articulation error probability $e(\text{SNR})$ could be thought of as being distributed over K -independent articulation bands. The bandwidth of each of these articulation bands was chosen so that they contribute equally to e [the articulation per critical band is constant from 0.3–7 kHz (Fletcher and Galt, 1950; Allen, 1994, 1996)]. Assuming band independence, the total articulation error may be written as a product over K band articulation errors

$$e = \epsilon_1 \epsilon_2 \cdots \epsilon_k \cdots \epsilon_K. \quad (13)$$

This equation is called the *band independence model*.

Galt established that the articulation bandwidth is proportional to cochlear critical bandwidths (French and Steinberg, 1947, page 93), as measured by the *critical ratio* method and the frequency jnd (Allen, 1994, 1996). Fletcher then estimated that each articulation band was the equivalent of 1 mm of distance along the basilar membrane, thereby taking up the 20-mm distance along the basilar membrane, between 300 to 8 kHz (Allen, 1996). Thus, the AI [Eq. (8)] may be viewed as an average SNR, *averaged over dB units*, of a scaled specific SNR, defined over cochlear critical bands.

As first derived in Allen (1994), the probability of articulation error in the k th band ϵ_k may be written in terms of the specific AI as

$$\epsilon_k = e^{\text{AI}_k / K}, \quad (14)$$

where the constant e_{\min} is defined as the minimum error via the relationship

$$e_{\min} \equiv 1 - \max_{\text{snr}}(P_c(\text{SNR})). \quad (15)$$

This constant e_{\min} depends in general on the corpus, talkers, and subjects. For Fletcher's work, e_{\min} was 1.5% ($\mathcal{H} \approx 11$, i.e., more than 2048 sounds). For the work reported here, a value of 0.254% ($\mathcal{H} = 4$) was used, based on an extrapolation of the MN55 data to $\text{AI} = 1$ and a minimization of the model parameters for a best fit to MN55 data.

It follows from the above relations that

$$P_c(\text{AI}) = 1 - e_{\min}^{\text{AI}}. \quad (16)$$

The total error $e = e_{\min}^{\text{AI}}$ in Eq. (16) was represented by Fletcher as $e = 10^{-\text{AI}/0.55}$. Both expressions are exponential in AI, differing only in the choice of the base ($e_{\min} = 10^{(-1/0.55)}$). Equation (16) only applies to the case of non-sense phones, having the maximum entropy.

Figure 5, left, shows the relative spectrum and noise level corresponding to SNRs of -18 to $+12$ dB, for female speech with a white-noise masker. On the right one may see the resulting $\text{AI}(\text{SNR})$, based on the calculations specified by the equations presented in this section. The final values of the AI were determined with $c = 2$ to be (starting from an SNR of $+12$): $[0.459, 0.306, 0.186, 0.1, 0.045, 0.016]$.

Because the spectrum of the speech and the spectrum of the noise are not the same, the $\text{AI}(\text{SNR})$ cannot be a linear function of SNR. Only for the case where the two spectra have the same shape will $\text{AI}(\text{SNR})$ be linear in SNR. For the case at hand, a white-noise masker, the high frequencies are progressively removed as the SNR decreases, as shown in the left panel of Fig. 5.

B. AM(SNR) to AM(AI)

The left panel of Fig. 6 shows the MN55 consonant identification curves $P_c^{(i)}(\text{SNR}) \equiv S_{i,i}(\text{SNR})$, as a function of the SNR for each of the 16 sounds ($i = 1, \dots, 16$), along with their mean $P_c(\text{SNR})$ (solid curve with circle symbols)

$$P_c \equiv \frac{1}{16} \sum_{i=1}^{16} P_c^{(i)}. \quad (17)$$

It must be mentioned that Eq. (17) only applies to the case at hand, where the *a priori* probabilities of the sounds are equal. In the more general case, a Bayesian formulation would be required.

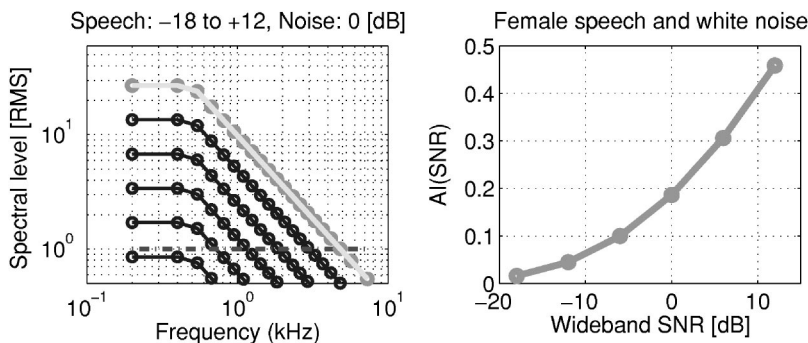


FIG. 5. Using the speech power spectrum given by the dashed line in Fig. 4, and assuming a uniform noise spectral level, the $\text{AI}(\text{SNR})$ was calculated. Each curve shows the relative spectral level of the speech having a peak level at the wideband SNRs used by Miller and Nicely $[-18, -12, -6, 0, 6, 12]$, in units of dB. The top curve shows the $+12$ -dB speech spectrum. The dashed-dot line is the noise spectral level having an rms of 0 dB.

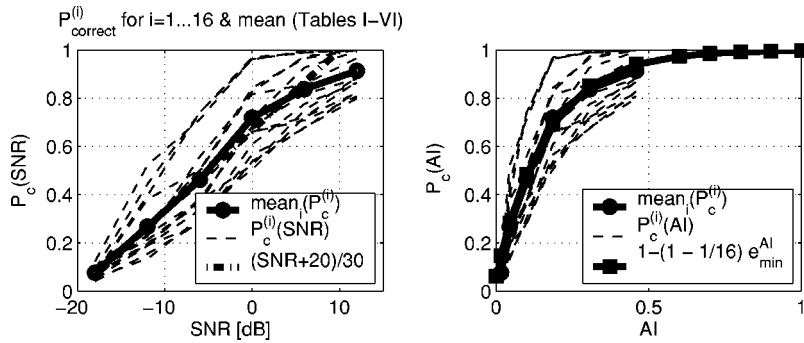


FIG. 6. The light dashed lines are $P_c^{(i)}$ for each of the 16 consonants. On the left the abscissa is the SNR in dB, while on the right, the AI is used as the independent variable. The solid-thick curve (circles) on both the left and right is the average score P_c , Eq. (17). The solid-thick curve (squares) on the right is the average phone prediction given by Eq. (18).

In the right panel the individual scores, along with the average, are shown as a function of the AI. To transform from SNR to AI the values shown in the right panel of Fig. 5 are used.

We also wish to compare the AI model prediction to the measurements shown in Fig. 6. However, it is necessary to modify Eq. (16) so that it accounts for chance (guessing) given by $P_{\text{chance}} = 2^{-\mathcal{H}}$, when $\mathcal{H}=4$ and $\text{AI}=0$. This is done by again assuming independence of the error probabilities. Since chance error for guessing is $e_{\text{chance}} = 1 - P_{\text{chance}}$, the chance-corrected $P_c(\text{AI})$ formula is

$$P_c(\text{AI}, \mathcal{H}) = 1 - e_{\text{chance}}(\mathcal{H})e_{\text{min}}^{\text{AI}}, \quad (18)$$

with

$$e_{\text{chance}}(\mathcal{H}) \equiv 1 - 2^{-\mathcal{H}}. \quad (19)$$

Fletcher's formula Eq. (16) is the limiting case of Eq. (18) when \mathcal{H} becomes large (Fletcher's $\mathcal{H} \approx 11$).

A plot of Eq. (18) is shown in the right panel of Fig. 6 (solid curve, square symbols), with $e_{\text{min}} = 0.254\%$, and $\mathcal{H}=4$. The fit of Eq. (18) to the average of the 16 MN55 curves is excellent.

Discussion

The left panel of Fig. 6 shows that there is an approximately linear relationship between $P_c(\text{SNR})$ and SNR over the range from -18 to 6 dB. The thick dashed-dot line is $(\text{SNR}+20)/30$. This line is useful as a simple reference.

The main deviation from the linear dash-dot curve is due to the strong saturation that occurs for the two nasal sounds and sound 7 [the three curves with the highest $P_c(\text{SNR})$]. Note that each of the sounds has a nearly linear $P_c^{(i)}(\text{SNR})$, with different saturation levels (if they are reached). The saturation point for $P_c(\text{SNR})$ occurs at an SNR of about 30 dB above the threshold, at -20 dB (thick solid line with circles). Note that since the relation $P_c(\text{SNR})$ depends on the noise spectrum, the linear relation observed in the left panel of Fig. 6 can only hold for the white-noise masker, since if the noise spectrum is changed, $P_c(\text{SNR})$ must change, and it is linear for the white-noise case.

In the right panel of Fig. 6 the extended AI model [Eq. (18)] is shown for MN55's data. Each of the 16 curves $P_c^{(i)}(\text{AI})$, $i=1, \dots, 16$, is shown as the light-dashed curves. This average [Eq. (17)] is shown as the solid-thick curve with circles.

The solid-thick line with squares is the extended (chance-corrected) AI model, Eq. (18). The value of e_{min} of

0.254% is one-sixth that used by Fletcher (1.5%). The smaller size could be attributed to the larger amount of training the subjects received over such a limited set size $\mathcal{H}=4 = \log_2(16)$.

As may be seen in the left panel of Fig. 5, since MN55 used white noise, the snr_k for frequency bands larger than about 0.7 kHz have an SNR of less than 30 dB, resulting in an AI of much less than 1. In fact, the AI was less than 0.5 for the MN55 experiment, corresponding to a maximum score of only 90%.

A most interesting and surprising finding is that the extended AI model [Eq. (18)] does a good job of fitting the average data. In fact, the accuracy of the fit over such a small set of just 16 consonants was totally unanticipated. This needs further elucidation.

C. Extended tests of the AI model

If one plots the total error probability $e(\text{AI}) = 1 - P_c(\text{AI})$ in log coordinates, as a function of AI, such plots should approximate straight lines. This follows from the log of Eq. (18)

$$\log(e(\text{AI})) = \log(e_{\text{min}})\text{AI} + \log(e_{\text{chance}}(\mathcal{H})), \quad (20)$$

which has the convenient form $y = ax + b$. The ordinate (y axis) intercept of these curves at $\text{AI}=0$ gives the log chance error [$b \equiv y(0) = \log(e(0)) = \log(e_{\text{chance}}(\mathcal{H}))$], while the ordinate intercept of these curves at $\text{AI}=1$ defines the sum of the log-chance error and the log-minimum error, namely [$a + b \equiv y(1)$, thus $a = \log(e_{\text{min}})$]. In Fig. 7 the log-error probabilities for each of the 16 sounds, along with the average and the AI model, are shown. The sounds have been regrouped so that the log-error plots have similar shapes. The shallow slopes are shown on the left and the steeper slopes on the right.

From Fig. 7, we shall find that the linear relationship [Eq. (20)] holds for 11 of the 16 sounds, with the free parameters $e_{\text{min}}(i)$ and $e_{\text{chance}}(i)$, either depending on the sound, or on a sound group.

The upper two panels show the most linear groups, while the lower panels are the most nonlinear (nonstraight) log-error curves. The curves that are close to linear (the two top panels) are consistent with the AI model, due to Eq. (20).

This observation of a log-linearity dependence for the probability of error of individual sounds is rather astounding in my view. First, there was no *a priori* basis for anticipating that individual sounds might obey Fletcher's band-independence property, Eq. (18). Second, if individual

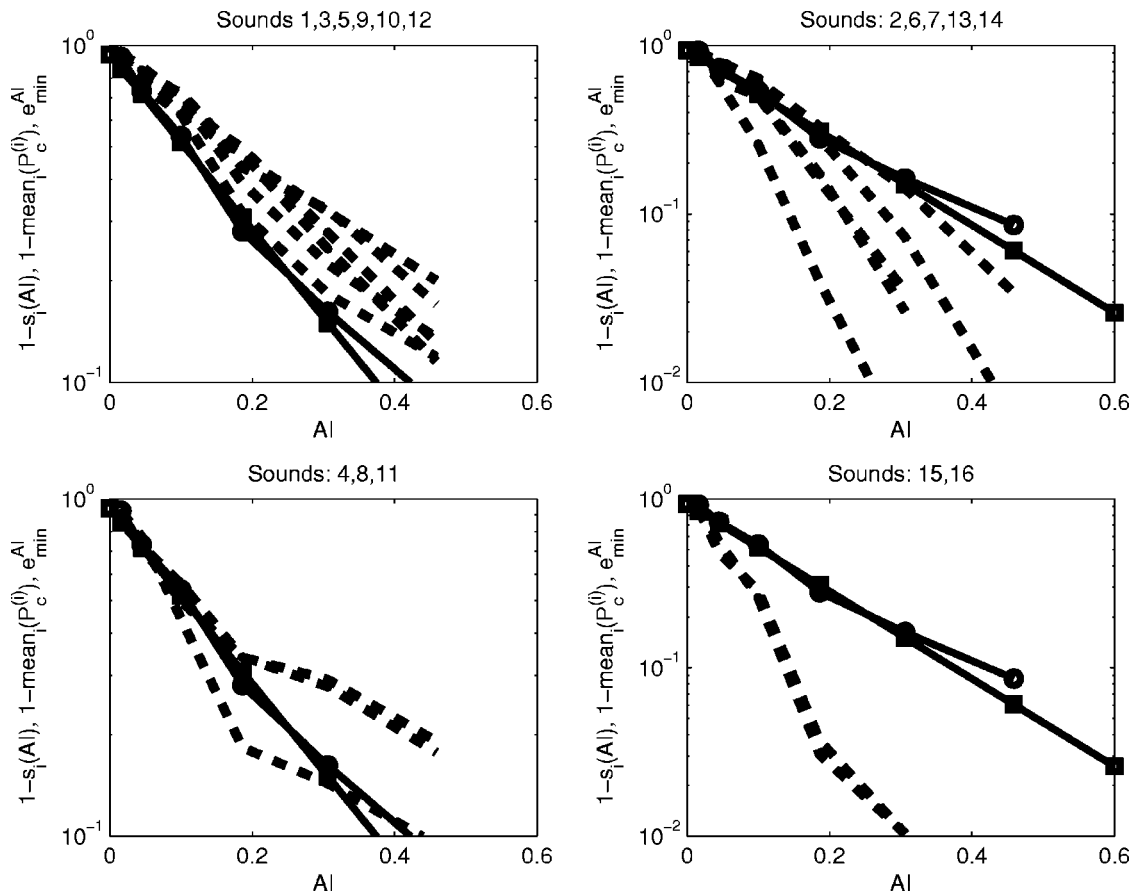


FIG. 7. This figure shows the probability of error for the i th sound, $P_e^{(i)}(AI) \equiv 1 - P_c^{(i)}(AI)$, as a dashed curve. To reduce the clutter, the sounds have been sorted over the four panels, with the sound number indicated in each panel title. The top two panels are the cases where the individual sound-error curves are close to straight lines. The left-upper panel are those cases where the sound lies above the average, while the right-upper panel shows those cases where the sound lies below the average. The two lower panels correspond to the sounds that violate the exponential rule (are not straight lines on a log-error plot). For reference, each panel contains the average probability of error $P_e(AI) \equiv 1 - P_c(AI)$, shown as the solid curve with circles, and the model error e_{\min}^{AI} , shown as the solid line (squares).

sounds obey equations of the form of Eq. (8), then sums of such equations cannot obey Eq. (8), since the sum of many exponentials, each having a different base, is not an exponential.

The finding that individual CV recognition error is exponential in the AI (the basis of the band-independence hypothesis) therefore extends, and at the same time violates, Fletcher's original fundamental AI hypothesis that the average error is exponential.

It is therefore essential to understand the source of the deviations for the individual sounds from the average, and to critically assess the accuracy of the model for individual sounds. Five sounds (4,8,11,15,16) have a probability of error that deviates from linear, with the most nonlinear and the largest deviations from the mean, being the nasals (15,16), as shown in the lower-right panel. In the next section I explore the reasons for this.

1. Log-error for the nasals

In Fig. 8 the nasal data are shown using the same log-error linear decomposition used in Fig. 3, where the total error (solid-thick curve) is the sum of the errors of the com-

peting sounds [i.e., Eq. (7)]. In the case of the nasals, the confusions for the other sounds is small, namely only /ma/ and /na/ significantly compete.

As a result of plotting the data as a function of AI, for $AI > AI_g = 0.045$ ($SNR \geq -12$), the log-error curves become linear in AI, as predicted (modeled) by Eq. (18). This value of AI_g is shown in the plot with an arrow indicating the point of separation of the target sound from the competing sound. Extrapolating this linear region back to $AI=0$, one finds the chance guessing probability of $1 - 2^{-\mathcal{H}_g} = 1/2$, corresponding to a nasal group entropy of $\mathcal{H}_g = 1$. This is shown on the graph by the dashed line superimposed on the corresponding error curve (stars). In the region $0 \leq AI \leq AI_g = 0.045$, \mathcal{H} depends on AI, since it dramatically drops from 4 to 1.

Thus, the reason that the nasal curves are not linear in Fig. 7 is that chance (the entropy factor) is dramatically changing between $0 \leq AI \leq AI_g$, due to the formation of the perceptual "event-nasal" group.

When the data are plotted as a function of SNR, as in Fig. 3, the log-error linearity is not observed. Also, the shape of the curve will depend on the spectrum of the noise. Clearly, the SNR to AI transformation is an important key to making sense of these data.

2. Log error for /pa/, /ta/, and /ka/

Finally, in Fig. 9 we return to the case of /pa/, /ta/, and /ka/. This 3-group generalizes the /ma/, /na/ 2-group conclusions of Fig. 8. In the middle panel it is clear that for small values of AI less than $0.045 S_{2,2}(\text{AI})$ for /ta/ is equal to the curves for /pa/ and /ka/ [$S_{2,j}(\text{AI}), j=1,3$]. As the AI rises above about 0.1, the three curves (circles) split due to the identification of /ta/ and the rejection of /pa/ and /ka/. The shape and slope of the curves corresponding to the two rejected sounds are identical. The projection of the rejected curves back to AI=0 gives the probability of chance error for a group of 3 (i.e., $1-1/3$), as shown by the dashed line in this middle panel. In the left-most and right-most panels, corresponding to /pa/ and /ka/, the two rejected sounds have very different log-error slopes. However, the two dashed curves still project back to the chance error probability for a group of 3 ($1-1/3$). This change in the slope for the two sounds shows that $e_{\min}(i,j)$ can, in general, depend on the sound in the group. This seems to reflect the more robust nature of /ta/ relative to /pa/ and /ka/ due to /ta/ having more high-frequency energy than its competitors.

Based on the small amount of the data shown in Fig. 8 and Fig. 9, it appears that the band-independence assumption [Eq. (13)] and the band error expression [Eq. (14)] model the individual sound confusions $S_{i,j}(\text{AI})$ more accurately than they model the average band error [Eq. (13)]. The total sound error is more precisely the sum of these off-diagonal confusion terms, as given by Eq. (7). The implications of this model seem quite significant, but without more data it is unwise to speculate further at this time.

IV. CONCLUSIONS

The intent of this paper is to provide a theoretical analysis of the venerable 1955 Miller and Nicely confusion matrix data, which have been difficult to fully appreciate, due to inadequate analysis methods. Replotting the data as a function of the SNR, rather than as confusions at a fixed SNR, provides a novel way of robustly clustering the feature groups. This grouping, not robustly defined in a single CM, is easily determined in such SNR plots. When working with individual CM data at a single SNR, clusters depend on the

sound ordering. When plotted as a function of SNR, sound order is irrelevant, and clusters depend instead on a smoothness, or continuity, across SNR.

A second contribution is the use of the AI as the independent variable. When the PI is plotted as a function of the SNR, the only structure observed are the clusters (Fig. 2). When these same data are plotted as a function of AI (Fig. 8), they become linear functions of AI (they form straight lines on log-error axes), consistent with the band-independence model of Fletcher [Eq. (13)], thereby corroborating Fletcher's AI model equation [Eq. (16)] for the case of single competing consonants. Plots of $S_{i,j}(\text{AI})$ depend on the spectrum of the noise.

A third contribution is the extension (and verification) of Fletcher's articulation model equation for $P_c(\text{AI})$, for the case of small set size [Eq. (18)], by introducing entropy \mathcal{H} into the model [Eq. (16)], thereby accounting for chance (guessing).

As the SNR increases from chance levels (e.g., -21 -dB SNR), sound groups form, forcing the entropy to decrease. The extended model [Eq. (18)] leads us to the conclusion that the entropy must depend on AI. This function, $\mathcal{H}(\text{AI})$, decreases from its maximum value of 4 at AI=0, to $\mathcal{H}_g \equiv \log_2(\text{group size})$ for AI=AI_g, where the group is fully formed. The entropy associated with these groups may be estimated from the clusters in $S_{ij}(\text{AI})$, or from the intercept at AI=0 of the dashed lines of Fig. 8 and Fig. 9. For the nasal sounds the group size is 2 ($\mathcal{H}=1$), leading to an intercept of $1-1/2$. For the [/pa/,/ta/,/ka/] group, the group size is 3 [$\mathcal{H}=\log_2(3)\approx 1.58$]; thus, the intercept is $1-1/3$.

A. Parametric model

In summary, a parametric model of the confusion matrices for the sound groups 1–3 and 15, 16 has been established. Chance, defined by $e_{\text{chance}}(\mathcal{H})$ [Eq. (19)], depends only on the experimental set (alphabet) size, characterized by $\mathcal{H}(0)$, not on the sounds themselves. Each sound may be described by three parameters. The sound-dependent parameters of Eq. (18) are $e_{\min}(i,j)$, $\mathcal{H}_g(i)$ and AI_g(i). The parameter $e_{\min}(i,j)$ appears to be a property of the individual

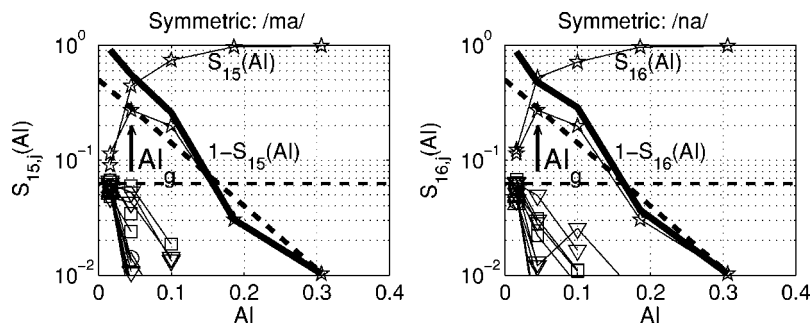


FIG. 8. Since the log-error plots for /ma/ and /na/ (see the lower-right panel of Fig. 7) show the greatest deviation from linear, they seem to be a “worst case” for the AI model. From this figure it is clear that the reason for the deviation from linear dependence is due to the migration of chance from $1/16$ ($\mathcal{H}=4$) to $1/2$ ($\mathcal{H}=1$), due to the nasal grouping. The rising nasal curves result from the robust grouping of the nasal, resulting in an increase in chance from $1/16$ at AI=0 to $1/2$ at AI ≈ 0.045 . The solid-thick curve is the sum of all the errors (and is $1 - P_c$ for the spoken sound). A dashed line has been drawn from the point (0, 0.5) to (0.31, 0.01). This line fits the error curve (/na/ given /ma/, and /ma/ given /na/) with very little error, for AI=AI_g>0.045, and intercepts the ordinate at $1/2$ for AI=0, as expected for a 2-group ($\mathcal{H}=1$). This further supports the band independence model Eq. (13).

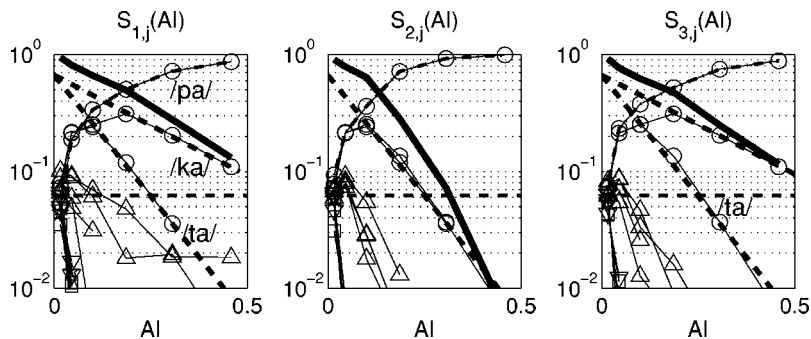


FIG. 9. This figure shows $S_{s,i}(AI)$ for $s=1, 2, 3$ corresponding to the sounds /pa/, /ta/, and /ka/. The dashed lines connect $(0, 1-1/3)$ with $(0.48, 0.1)$ and $(0.442, 0.01)$.

sounds in the group. For example, the value of $e_{\min}(i,j)$ for sounds 1 and 3 are the same, but the value for sound 2 is much smaller (Fig. 9).

The value of e_i is the sum of all the errors of the corresponding off-diagonal values, namely

$$e_i(AI) = \sum_{j \neq i} S_{i,j}(AI). \quad (21)$$

This equation is the same as Eq. (7) except for the independent variable. The total minimum error is given by the average of the row errors, evaluated at $AI=1$

$$e_{\min} = \frac{1}{16} \sum_i e_i(AI) \Big|_{AI=1}. \quad (22)$$

This follows from Eq. (17) and Eq. (7).

Parameter $AI_g(i)$ characterizes the transition from the maximum alphabet entropy [$\mathcal{H}(AI=0)$] to the group entropy \mathcal{H}_g . We have not attempted to find an analytical expression for $H(AI)$, to describe the transition from entropy maximum $\mathcal{H}(0)$ to that of the group $\mathcal{H}(AI_g)$. The hierarchies mentioned in the Introduction each have their own AI_g parameter, each group within the hierarchy having a smaller value of AI_g .

The change in the entropy, from $\mathcal{H} \rightarrow \mathcal{H}_g$, due to the formation of groups as the sounds start becoming identified, accounts for the deviations from linear of the log-error probability.

One might even view the ultimate identification of the sound, as $AI \rightarrow 1$, as a further reduction of the sound's row entropy to zero.

Since the parametric model is based on very little data, there is presently no clue as to how it will generalize.

B. Preprocessing of \mathcal{C}

The confusion matrix \mathcal{C} was transformed to form the articulation matrix \mathcal{A} , by a normalization step [Eq. (1)], and then further transformed by symmetrizing [Eq. (3)]. These two transformations interact and must therefore be iterated to convergence. While these transformation steps are not essential, they are justified, since they reduce sampling noise and remove subject bias. Sampling noise and bias are interesting topics in their own right that deserve further analysis. For example, the skew-symmetric form $A(AI, \mathcal{H})$ [Eq. (4)] should be carefully considered in the future, to characterize and determine the precise nature of the subject bias. This bias should be considered when designing MN55-type experiments.

Once the CM has been preprocessed so that its rows sum to 1, an error decomposition is possible. The equation for this is Eq. (7), which says that the total sound identification error is the sum of the confusions. This expression is useful in those cases where the sounds group, as it uniquely decomposes the error into the group confusions. When plotted as log error (e.g., Fig. 9) one may characterize the sources of the sound errors in a quantitative way [i.e., using the parameters $e_{\min}(i,j)$ and $AI_g(i)$]. This method seems superior to all previous analysis methods of such confusion matrices.

As shown in Fig. 7, the nasals appear to violate Fletcher's independence formula Eq. (13), since the total error is not a straight line on the log-error plot. However, when decomposed by Eq. (7), we see that *individual competing sounds obey band independence*. Thus the total error deviates from a linear log error, due to the dramatic change in \mathcal{H} with AI , from 4 to 1, over the small range of AI between 0 and $AI_g \approx 0.045$ (Fig. 8).

Furthermore, the projection of the straight lines that lie along the log-error curves, back to $AI=0$, gives e_{chance} for the group. This is an important corroboration of Eq. (18). Two examples of these are seen for the nasals, where the group has entropy 1, and the dashed lines of Fig. 8 project back to $1-1/2$, and the 3-group of Fig. 9, where the dashed lines project back to $1-1/3$.

C. Calculating the AI

The procedure for calculating the AI, developed in Sec. III A, has some novel aspects as well. Rather than defining the specific AI in terms of the band SNRs, the modified function of Eq. (11) was used. The justification for Eq. (11) comes from the work of Fletcher as well as French and Steinberg, both of whom promote (but did not use) this detection formulation. The *speech detection constant* c is chosen to characterize the detection of the speech peaks when noise is added to the speech. Even though this formulation of the AI has some important advantages, and is more accurate, it is never referred to in the modern AI literature. This, I feel, is a mistake that needs rectification. Again, this approach was not studied in detail in this paper; however, there is a detailed analysis of Eq. (11) in both of the references, and a deeper analysis here is off topic. It was necessary to introduce Eq. (11) to get reasonable values of $AI(\text{SNR})$ when fitting the model $P_c(AI, \mathcal{H})$, as shown in Fig. 6. This is because the estimates of the SNR as a function of frequency [i.e., $\text{snr}_k(\text{SNR})$], in the left panel of Fig. 5, are strongly affected by this detection model, and on the specific choice of c in

Eq. (10). Without the use of this speech detection parameter ($c=2$), the band SNR values $\text{snr}_k(\text{SNR})$ would be unrealistic for small values of AI.

D. Band independence

Fletcher's band-independence assumption Eq. (13) has proven to be an important tool at the individual sound level. This should come as a surprise, as it was not anticipated by Fletcher's work, or any work following (that I am aware of). On the other hand, the fact that it works at all should lead us to the possibility that it could generalize. It would appear from the analysis provided here that Eq. (13) is more accurate in describing competing sounds than in describing the average probability correct $P_c(\text{AI}, \mathcal{H})$. This statement is supported by the very linear behavior of the off-diagonal confusion terms $S_{i,j}$ in Fig. 8 and Fig. 9. The partial errors are highly linear once the group has formed ($\text{AI} > \text{AI}_g$). It follows from Eq. (7) that the deviations from linear are a result of the groups, which depend on the noise spectrum. The influence of a group formation distorts this basic linear character, and therefore distorts the linearity of the sum over many error terms [i.e., $P_c(\text{AI})$]. Based on the small amount of data we presently have (those shown in this paper), it would be reasonable to conclude that band independence is more a property of individual consonants than it is of the group means, as first proposed (derived) by Fletcher. Much more data and analysis are needed to verify this possibility, which is hardly proved at this time.

E. Implications to ASR

Automatic computer recognition of speech (ASR) could benefit from many of the same considerations as those of MN55. It would be interesting to run similar experiments on modern ASR systems, to characterize their $\mathcal{A}(\text{AI}, \mathcal{H})$ performance. In many cases this might not be practical, due to the limited performance of the ASR front ends, or if the confusion matrices turned out to be skew-symmetric. The ASR language model performance is inhibited when using nonsense speech, since most of these systems depend on some sort of language context for their performance.

ACKNOWLEDGMENTS

The inspiration for this work started with a question by David Nahamoo which I could not answer: "What is the

meaning of e_{\min} ?" I would like to thank my students Suvrat Budhlakoti, Bryce Lobdell, Andrew Lovitt, and especially Sandeep Phatak, and also thank Harry Levitt, for many important insights and critical discussion, and thank Anthony Watkins and two anonymous reviewers, for many insightful comments. Finally, I would especially like to thank George Miller for doing his original 1955 work, for reading the present manuscript, and for the personal encouragement he has provided.

- Allen, J. B. (1994). "How do humans process and recognize speech?" *IEEE Trans. Speech Audio Process.* **2**(4), 567–577.
- Allen, J. B. (1996). "Harvey Fletcher's role in the creation of communication acoustics." *J. Acoust. Soc. Am.* **99**(4), 1825–1839.
- ANSI S3.5-1997 (1997). "Methods for calculation of the speech intelligibility index (SII-97)" (American National Standards Institute, New York).
- Fletcher, H. (1995). "Speech and hearing in communication," in *The ASA Edition of Speech and Hearing in Communication*, edited by J. B. Allen (Acoustical Society of America, New York).
- Fletcher, H., and Galt, R. (1950). "Perception of speech and its relation to telephony," *J. Acoust. Soc. Am.* **22**, 89–151.
- Fletcher, H., and Munson, W. (1937). "Relation between loudness and masking," *J. Acoust. Soc. Am.* **9**, 1–10.
- French, N. R., and Steinberg, J. C. (1947). "Factors governing the intelligibility of speech sounds," *J. Acoust. Soc. Am.* **19**, 90–119.
- Goldstein, L. (1980). "Bias and asymmetry in speech perception," in *Errors in Linguistic Performance*, edited by V. A. Fromkin (Academic, New York), Chap. 17, pp. 241–261.
- Houtgast, T., and Steeneken, H. (1973). "The modulation transfer function in room acoustics as a predictor of speech intelligibility," *Acustica* **28**, 66–73.
- Kryter, K. D. (1962a). "Methods for the calculation and use of the articulation index," *J. Acoust. Soc. Am.* **34**(11), 1689–1697.
- Kryter, K. D. (1962b). "Validation of the articulation index," *J. Acoust. Soc. Am.* **34**(11), 1698–1702.
- Miller, G. A. (1951). *Language and Communication* (McGraw Hill, New York).
- Miller, G. A., and Nicely, P. E. (1955). "An analysis of perceptual confusions among some English consonants," *J. Acoust. Soc. Am.* **27**(2), 338–352.
- Miller, G. A., Heise, G. A., and Lichten, W. (1951). "The intelligibility of speech as a function of the context of the test material," *J. Exp. Psychol.* **41**, 329–335.
- Steeneken, H., and Houtgast, T. (1980). "A physical method for measuring speech-transmission quality," *J. Acoust. Soc. Am.* **67**(1), 318–326.
- Wang, M. D., and Bilger, R. C. (1973). "Consonant confusions in noise: A study of perceptual features," *J. Acoust. Soc. Am.* **54**, 1248–1266.

Coherence and the speech intelligibility index

James M. Kates^{a)}

*GN ReSound and University of Colorado, Department of Speech Language and Hearing Sciences,
409 UCB, Boulder, Colorado 80309*

Kathryn H. Arehart^{b)}

*University of Colorado, Department of Speech Language and Hearing Sciences, 409 UCB,
Boulder, Colorado 80309*

(Received 15 July 2004; revised 6 January 2005; accepted 10 January 2005)

The speech intelligibility index (SII) (ANSI S3.5-1997) provides a means for estimating speech intelligibility under conditions of additive stationary noise or bandwidth reduction. The SII concept for estimating intelligibility is extended in this paper to include broadband peak-clipping and center-clipping distortion, with the coherence between the input and output signals used to estimate the noise and distortion effects. The speech intelligibility predictions using the new procedure are compared with intelligibility scores obtained from normal-hearing and hearing-impaired subjects for conditions of additive noise and peak-clipping and center-clipping distortion. The most effective procedure divides the speech signal into low-, mid-, and high-level regions, computes the coherence SII separately for the signal segments in each region, and then estimates intelligibility from a weighted combination of the three coherence SII values. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1862575]

PACS numbers: 43.71.Gv, 43.72.Dv, 43.71.Ky, 43.71.An [PFA]

Pages: 2224–2237

I. INTRODUCTION

A hearing aid is intended to improve speech intelligibility for the hearing-impaired user, but it is often difficult to predict the amount of benefit that will be obtained in practice. Intelligibility is reduced not only by the hearing loss, but also by ambient noise, room reverberation, and the noise and nonlinear distortion introduced by the hearing aid. The desire to assess speech intelligibility under adverse conditions has led to the development of several measurement and prediction procedures. Some of the problems affecting speech intelligibility, such as additive noise and bandwidth reduction, have been addressed by the speech intelligibility index (SII) (ANSI S3.5-1997), and reverberation has been addressed by the speech transmission index (STI) (Steeneken and Houtgast, 1980; Houtgast and Steeneken, 1985). However, there is no metric that can successfully deal both with noise and with the various types of nonlinear distortion that can occur in a hearing aid or communication system.

An intelligibility-estimation procedure for hearing aids must satisfy several criteria. First, the metric must be applicable to systems, like hearing aids, that have frequency-dependent magnitude and phase transfer functions. A flat frequency response, often assumed in telephony and for digital speech coding systems, cannot be assumed for a hearing aid. Second, in the presence of additive noise without distortion, the metric should yield intelligibility predictions that are consistent with those produced by the SII calculation (ANSI S3.5-1997). Third, the method should be applicable to speech as the test signal; much of the nonlinear signal processing in hearing aids is specifically designed for speech as the input and will respond differently to other types of exci-

tation such as pure tones, multitone complexes, or broadband noise.

The goal of the work reported in this paper is to derive a speech intelligibility metric that generalizes the SII. The specific objectives of the work are twofold: first, the metric should accurately predict the effects of broadband noise and distortion for speech reproduced by hearing aids and other communication systems, and, second, the metric should be accurate for hearing-impaired as well as normal-hearing listeners.

A. Speech intelligibility in normal-hearing listeners

There are several procedures for estimating speech intelligibility in normal-hearing listeners. The best known are the articulation index (AI) (French and Steinberg, 1947; Kryter, 1962), which has been further developed to produce the speech intelligibility index (ANSI S3.5-1997), and the speech transmission index (Steeneken and Houtgast, 1980; Houtgast and Steeneken, 1985). The SII measures the signal-to-noise ratio (SNR) on a dB scale in each frequency band. The calculation procedure then adjusts for auditory threshold and for frequency-domain masking effects and sums the weighted SNR across frequency to produce the intelligibility estimate. The STI uses amplitude-modulated speech-shaped noise as the excitation. Its calculation procedure measures the modulation depth in each frequency band in the listening environment and forms a weighted sum across frequency to give an estimate of the SNR. The STI calculation can also be performed using speech as the stimulus (Payton and Braida, 1999). The SII is effective in stationary additive noise and for filtering that restricts the speech bandwidth, and the STI is also effective in reverberation.

The SII procedure embodies a simplified model of the auditory periphery. The critical band formulation of the SII

^{a)}Electronic mail: jkates@gnresound.dk

^{b)}Electronic mail: kathryn.arehart@colorado.edu

procedure given in Table I of the standard (ANSI S3.5-1997) uses the critical bands of Zwicker (1961), which give a representation of auditory frequency analysis. The auditory threshold for normal and impaired hearing is incorporated into the calculation in Eq. (10) of the standard, and upward spread of masking for normal ears is included in Eqs. (6)–(9) of the standard. Because of this implicit auditory model, the SII procedure can be extended to include conditions not incorporated into the standard as long as the basic physical and perceptual assumptions are not violated.

Examples of previous extensions to the SII are the work of Kates (1987) and Rhebergen and Versfeld (2004). To compute the SII, the speech power spectrum and the noise power spectrum are averaged across the entire duration of the stimulus, and the SII is then calculated using these long-term averages. The SII computed this way is accurate for stationary noise, but is inaccurate for fluctuating noise. Kates (1987) evaluated an adaptive noise-cancellation system by dividing the speech into segments, computing the AI for each segment, and then plotting the change in the AI as a function of time to illustrate the system convergence. Rhebergen and Versfeld (2004) have extended the SII approach to predict the effects of fluctuating noise. They propose dividing the noisy speech into segments, computing the SII for each segment, and then averaging the SII values. The SII assumptions concerning auditory frequency resolution, threshold, and masking are preserved in their procedure, but the analysis adds additional understanding of the effects of signal fluctuations and the ability to listen in the noise valleys.

The principles underlying the SII have also been applied to estimating speech intelligibility and quality for distortion mechanisms such as symmetric peak clipping. Steeneken and Houtgast (1980), for example, show that the modulation transfer function can accurately predict speech intelligibility for peak-clipped speech. Kates and Kozma-Spytek (1994) studied speech subjected to low-frequency or high-frequency amplification followed by symmetric peak clipping. They computed the signal-to-distortion ratio (SDR) from the magnitude-squared coherence (MSC), and then derived a speech-quality metric based on a weighted sum of the SDR across frequency. This metric, which is similar to the frequency-dependent weighting of the SNR used in the SII calculation, was accurate in predicting speech quality for peak clipping even though auditory masking was not included. An extension of this approach (Kozma-Spytek *et al.*, 1996) was accurate in predicting speech quality for peak-clipped speech for hearing-impaired listeners, although some of the hearing-impaired subjects appeared to concentrate more on the low frequencies and other subjects appeared to concentrate more on the high frequencies in making their judgments.

B. Speech intelligibility in hearing-impaired listeners

Speech intelligibility in hearing-impaired listeners has been studied for conditions of bandwidth reduction and additive noise, but only rarely for distortion. For additive noise and bandwidth reduction, the SII (ANSI S3.5-1997) procedure has been found in general to overestimate speech intelligibility for hearing-impaired listeners (Pavlovic, 1984;

Ludvigsen, 1987; Hornsby and Ricketts, 2003). One solution is to apply a desensitization factor to the SII calculation, reducing the estimated intelligibility in proportion to the measured hearing loss (Pavlovic *et al.*, 1986; Magnusson *et al.*, 2001). The desensitization factor does not explain the reason for the discrepancy between the SII and the observed intelligibility, but it does provide an accurate empirical correction. The desensitization factor accuracy can be further improved by applying individually determined frequency-dependent proficiency factors in the SII calculation (Ching *et al.*, 1998).

A second solution to the problem of overestimating the intelligibility for hearing-impaired listeners is to modify the effective signal-to-noise ratio (SNR) used in the SII calculation. For example, the upward spread of masking function used in the SII calculation is based on averaged measurements of normal-hearing subjects. Increasing the spread of masking for the impaired ear reduces the effective signal-to-noise ratio (SNR) above the frequency region of the masker. The reduction in SNR reduces the calculated SII value. Rankovic (1998) used individual frequency-domain masking functions in applying the Fletcher and Galt (1950) calculation procedure to hearing-impaired listeners and found that the individualized calculations improved the accuracy of the intelligibility predictions without needing a desensitization factor. Hornsby and Ricketts (2003) measured pure-tone detection thresholds in speech-shaped noise for each hearing-impaired listener and used these data to increase the effective masking spectrum in the SII calculation. This modification of the SII implicitly includes the increase in spread of masking in the impaired ear and thus reduces the effective SNR. The modification resulted in intelligibility predictions for hearing-impaired subjects, having flat losses, that were comparable in accuracy to those for normal-hearing listeners, again without the need for a desensitization factor.

Crain and Van Tasell (1994) studied speech intelligibility in hearing-impaired subjects for peak-clipping distortion. They found that peak-clipping distortion increased the speech reception threshold (SRT) more for hearing-impaired listeners than for normal-hearing listeners. Crain and Van Tasell compared the SRT for no distortion to the SRT for infinite peak clipping after equating the levels of presentation. They found that the SRT for the hearing-impaired subjects shifted by 5 dB while that for the normal-hearing subjects shifted by only 2 dB. The shift in SRT was also found to be highly correlated with the reduction in the coherence between the input and distorted signals.

C. Test signal

Speech is the desired test signal for most types of hearing-aid processing. In many nonlinear processing systems, such as the noise suppression and multiband compression found in digital hearing aids, the signal processing behavior depends on the characteristics of the signal. For example, replacing speech with the amplitude-modulated speech-shaped noise used in the STI measurement (Steeneken and Houtgast, 1980) or with a synthesized approximation to speech (CCITT, 1986) will result in different frequency- and time-dependent gains than would occur for

real speech. Thus estimating the distortion that will occur for speech in a nonlinear speech-processing system requires the use of speech or, equivalently, a speechlike test signal that has statistics very close to those of real speech.

D. Coherence measurements and hearing aids

Coherence measurements are becoming important in evaluating hearing aids. Coherence measurement procedures using a broadband noise input signal have been developed for hearing aids (Preves *et al.*, 1989; Dyrland, 1989; Preves, 1990; Kates, 1992; Fortune and Preves, 1992; Schneider and Jamieson, 1995; Kates, 2000), and measurement procedures were standardized in ANSI S3.42-1992. The use of a broadband stimulus is essential for determining the response of multi-channel compression hearing aids; in these systems the gain as a function of frequency depends on both the signal amplitude and spectral distribution. Coherence is attractive for determining the nonlinear behavior of such systems because it measures all forms of distortion and not just the harmonic distortion traditionally measured (ANSI S3.22-1996).

An implicit assumption in using coherence to measure the nonlinear system behavior is that equal reductions in coherence will cause equal reductions in intelligibility or sound quality. In dealing with a single distortion mechanism, such as amplifier saturation, a correlation between a reduction in sound quality and a reduction in coherence is often found (Palmer *et al.*, 1995). However, different distortion mechanisms may reduce the MSC by the same amount yet have different perceptual consequences. The relationship between coherence and intelligibility and sound quality has not been investigated for different forms of distortion, and the assumption that equal reductions in MSC are perceptually equivalent has not been tested.

E. Summary of the paper

In this paper the SII approach is extended to estimate intelligibility for peak clipping and center clipping distortion as well as for additive noise. The new procedure is a modified form of the SII procedure wherein the standard speech signal-to-noise ratio estimate is replaced by the signal-to-distortion ratio computed from the MSC. The MSC is used because it is immune to linear transformations of the signal, is reduced by nonlinear distortion as well as by additive noise, and because it is applicable to any excitation signal, including speech. The SII is used as the basis of the metric because of its established accuracy in predicting intelligibility for additive noise and because it includes a model of auditory masking as well as the threshold shift in impaired ears. The new metric is tested with both normal-hearing and hearing-impaired listeners.

The remainder of this paper describes the evaluation and modeling of speech intelligibility scores from normal-hearing and hearing-impaired individuals. The stimuli used in the evaluations are sentences subjected to different forms of noise and distortion. The first experiment uses normal-hearing listeners. The subject intelligibility scores are modeled using a modification of the SII procedure in which co-

herence is used to estimate the effects of the noise and distortion on the speech. It is shown that the predictions using the coherence SII do not adequately model the intelligibility results. A further modification, based on computing the coherence SII separately for low-, medium-, and high-level segments of each sentence, is shown to be much more accurate in predicting the intelligibility scores.

The second experiment uses hearing-impaired listeners. The processed sentence stimuli are the same as used for the normal-hearing subjects. However, the hearing-impaired subjects are provided with amplification to compensate for each individual hearing loss. The intelligibility scores for the hearing-impaired subjects are predicted using the three-level coherence SII procedure with the same weights for the low-, mid-, and high-level segments of the sentences as used for the normal-hearing subjects. The only modification is the inclusion of the individual hearing loss in the coherence SII calculations. The new procedure is shown to be just as accurate in predicting the intelligibility for the hearing-impaired listeners as it is for the normal-hearing listeners.

II. EXPERIMENT I: NORMAL-HEARING LISTENERS

A. Listeners

Thirteen adult listeners with normal hearing (age range 18 to 45 years) participated. All listeners had thresholds of 20 dB HL (ANSI, 1989) or better at octave frequencies from 250 to 8000 Hz, inclusive. Listeners were tested individually in a double-walled sound booth and were compensated \$10/hour for their participation.

B. Test materials

The test materials consisted of the Hearing-in-Noise-Test (HINT) sentences (Nilsson *et al.*, 1994). The sentences were digitized at a 44.1-kHz sampling rate and down-sampled to 22.05 kHz to reduce computation time. There are 25 lists with ten sentences per list, giving a total of 250 sentences. Each test sentence was combined with additive noise or was subjected to symmetric peak-clipping distortion or symmetric center-clipping distortion. The two forms of clipping were chosen as examples of memoryless nonlinearities that have been used in previous investigations (Licklider, 1946). Peak clipping is related to arithmetic, amplifier, and receiver saturation in a hearing aid. Center clipping is related to noise-suppression systems that reduce the amplitude of low-level portions of the signal.

The additive noise was extracted from the opposite channel of the HINT test compact disc. The noise has the same long-term spectrum as the sentences. The signal-to-noise ratio (SNR) was determined by computing the root-mean-squared (RMS) power of each sentence, ignoring the silent intervals at the beginning and end of the sentence, and adjusting the noise power over the same interval to give the desired SNR. The combined signal-plus-noise power over the duration of each sentence was then adjusted to give a presentation level of 65 dB SPL.

The peak-clipping and center-clipping distortion thresholds were set using the histogram of the magnitudes of the signal samples for each sentence. The silent intervals at the

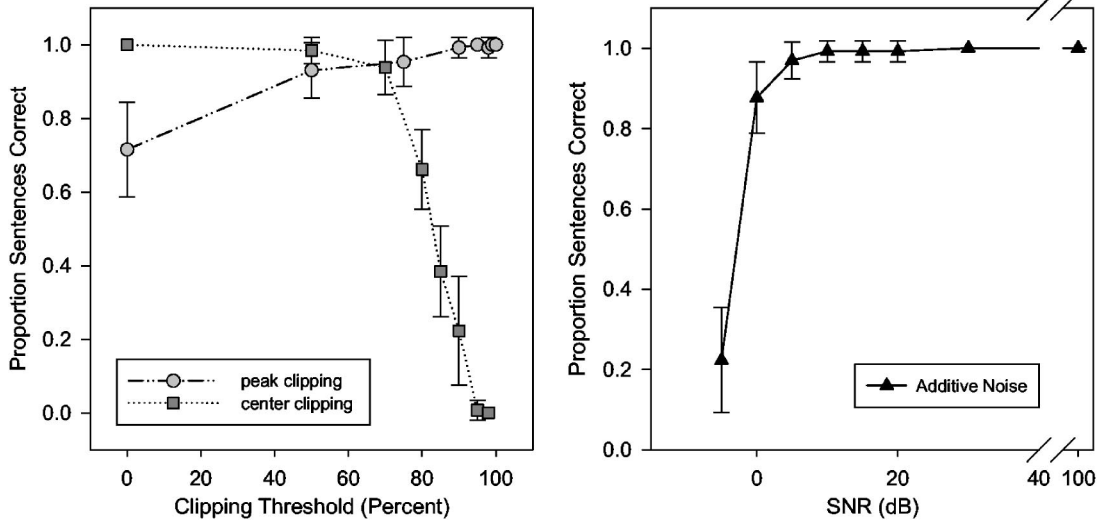


FIG. 1. Average subject intelligibility scores for additive speech-shaped noise (right panel) and peak-clipping and center-clipping distortion (left panel) for the HINT sentence test materials. The error bars show one standard deviation.

beginning and end of each sentence were discarded, and the cumulative distribution of the absolute values of the signal samples was then computed for the sentence. The clipping threshold was set as a percent of the cumulative magnitude histogram for the sentence. For symmetric peak clipping, the clipping operation is given by

$$y(n) = \begin{cases} c, & x(n) > c, \\ x(n), & -c \leq x(n) \leq c, \\ -c, & x(n) < -c \end{cases} \quad (1)$$

where $x(n)$ is the speech input, $y(n)$ is the distorted output, and c is the clipping threshold. The symmetric center clipping operation is given by

$$y(n) = \begin{cases} x(n), & x(n) > c, \\ 0, & -c \leq x(n) \leq c, \\ x(n), & x(n) < -c \end{cases} \quad (2)$$

The distorted sentences were then readjusted to give an average level of 65 dB SPL.

For each individual listener, 24 lists with ten sentences per list were selected for processing, with eight lists randomly selected for each of the three distortion conditions. Eight levels of distortion were evaluated for each distortion type and the distortion levels were randomized within each block. The distortion levels were peak clipping with the clipping threshold set to {0%, 50%, 75%, 90%, 95%, 98%, 99%, 100%} of the cumulative magnitude histogram of each sentence; center clipping with the clipping threshold set to {0%, 50%, 70%, 80%, 85%, 90%, 95%, 98%} of the cumulative magnitude histogram; and additive noise with the SNR set to {-5, 0, 5, 10, 15, 20, 30, 100} dB. Note that the effects of peak clipping are reduced as the clipping threshold is increased, while the effects of center clipping are increased as the clipping threshold is increased.

C. Procedure

Listeners participated in 1-h sessions. During the session, the listeners were presented with three blocks of 80 sentences, with one block in each of the three distortion conditions. The order of presentation of the three distortion conditions was randomized across listeners and across sessions. The order of presentation of sentences within a block was also randomized for each listener. During the sessions, listeners verbally repeated each sentence after it was presented. The tester then scored the proportion of complete HINT sentences that were correctly repeated by the listener.

For listener presentation, the digitally stored speech stimuli went through a digital-to-analog converter (Tucker-Davis Technologies RP2), an attenuator (TDT PA5), and a headphone buffer amplifier (TDT HB7). Finally, the stimuli were presented monaurally to the left ear of each listener through a TDH-49 earphone. All stimuli were presented to the listeners at an equalized-RMS level of 65 dB SPL.

D. Intelligibility scores

The intelligibility scores for the normal-hearing subjects are plotted in Fig. 1. The scores for each listener were first averaged across the ten sentences used for each combination of distortion type and level. The average intelligibility scores for each condition were then averaged across the 13 listeners. The average intelligibility (proportion sentences correct) is plotted as a function of SNR in dB (right panel) and as a function of the clipping threshold expressed as a percent of the cumulative amplitude distribution (left panel). The error bars show the standard deviation across the 13 subjects.

For the additive noise, the intelligibility increases as the SNR increases. The general shape of the curve is consistent with previous intelligibility results for the HINT sentences in additive speech-shaped noise (Nilsson *et al.*, 1994; Eisenberg *et al.*, 1998). The speech intelligibility is a steep function of SNR over the range of -5 to +5 dB; the intelligibility goes from 22% to 97%, giving an average slope of

7.5% per dB. The intelligibility reaches an asymptote of nearly perfect sentence recognition above the 5-dB SNR.

The worst case intelligibility for peak clipping is 72% correct, obtained for infinite clipping (replacing the speech signal samples by their signs). The intelligibility improves as the clipping threshold increases. The high intelligibility for peak clipping indicates that much of the vowel and consonant structure of the speech is preserved even for infinite clipping. For center clipping, the intelligibility remains high for a clipping threshold up to 70%, and then decreases rapidly with increasing clipping threshold. At a clipping threshold of 98% all but the highest peaks of the speech signal have been removed, resulting in an intelligibility of 0%. These results are consistent with previous results for peak-clipped and center-clipped speech (Licklider, 1946).

III. COHERENCE

The coherence function (Carter *et al.*, 1973) is the normalized cross-spectral density given by

$$\gamma(\omega) = S_{xy}(\omega) / [S_{xx}(\omega)S_{yy}(\omega)]^{1/2}, \quad (3)$$

where $S_{xy}(\omega)$ is the cross-spectral density between the two stationary random processes $x(n)$ and $y(n)$ having autospectral densities $S_{xx}(\omega)$ and $S_{yy}(\omega)$, respectively. The magnitude-squared coherence function (MSC) is then given by

$$|\gamma(\omega)|^2 = |S_{xy}(\omega)|^2 / S_{xx}(\omega)S_{yy}(\omega), \quad (4)$$

with $0 \leq |\gamma(\omega)|^2 \leq 1$.

When $x(n)$ is the input to a system and $y(n)$ is the output, the MSC represents the fraction of the output signal power that is linearly dependent on the input at each analysis frequency. The complementary fraction $1 - |\gamma(\omega)|^2$ gives the output power at each frequency that is unrelated to the input at that frequency and thus represents the nonlinear distortion and noise. The signal-to-noise and distortion ratio (SDR) is therefore given by (Carter *et al.*, 1973)

$$\text{SDR}(\omega) = |\gamma(\omega)|^2 / [1 - |\gamma(\omega)|^2]. \quad (5)$$

When $|\gamma(\omega)|^2 = 1$, all of the output is linearly related to the input and the SDR approaches infinity. When $|\gamma(\omega)|^2 = 0$, there is no relationship between the output and the input and the SDR is zero as well.

The MSC is typically estimated using fast Fourier transform (FFT) techniques (Carter *et al.*, 1973). The data sequences $x(n)$ and $y(n)$ are divided into a number M of overlapping windowed data segments. The cross-spectrum and autospectra are computed for each segment using the FFT and are then averaged across segments, and the MSC is computed from the averages. For M data segments, the estimated MSC is given by

$$|\gamma(k)|^2 = \frac{|\sum_{m=0}^{M-1} X_m(k) Y_m^*(k)|^2}{\sum_{m=0}^{M-1} |X_m(k)|^2 \sum_{m=0}^{M-1} |Y_m(k)|^2}, \quad (6)$$

where the asterisk denotes the complex conjugate. $X_m(k)$ and $Y_m(k)$ are the spectra of the m th windowed data segments of $x(n)$ and $y(n)$, respectively, computed using the FFT algorithm where k is the FFT bin index. Typically, a

Hann or Hamming data window and 50% overlap are used for the data segments. The estimated MSC is subject to bias effects due to the finite number of segments used in the FFT procedure and to the interaction of the processing system delay with the segment length. For the greatest accuracy, unbiasing techniques that temporally align the input and output sequences should be used when computing the MSC (Kates, 1992).

The coherence is unaffected by a linear transformation of the signal. A linear filter will increase the gain at each frequency, but the change in gain affects the numerator and denominator of Eq. (6) by equal amounts. Nonlinear transformations, however, will reduce the MSC (Kates, 2000). Thus the MSC is reduced by additive noise, but it is also reduced by nonlinear distortion such as peak clipping, center clipping, and by dynamic-range compression.

IV. THE COHERENCE SII

A. Calculation procedure

The SII standard (ANSI S3.5-1997) allows the SNR to be calculated using the speech and noise spectra measured in octave, one-third octave, or auditory critical bands. The critical-band SII procedure is used in this study, and the filter center frequencies and bandwidths are given in Table I of ANSI S3.5-1997. The critical band filter shapes used in this study are the simplified ro-ex filters suggested by Moore and Glasberg (1983). Let q_j be the center frequency in kHz of the j th critical band, with the band center frequency shifted to match the center frequency of the closest FFT analysis bin. The filter bandwidth from the ANSI Table I is denoted by b_j ; the bandwidths correspond to the classical critical bands (Zwicker and Terhardt, 1980). The ro-ex filter parameter is given by

$$p_j = \frac{4(1000q_j)}{b_j}, \quad (7)$$

where the factor of 1000 converts the filter center frequency to Hz. The simplified ro-ex filter is then given by

$$W_j(f) = (1 + p_j g) \exp(-p_j g), \quad (8)$$

where g is the normalized deviation from the filter center frequency given by

$$g = |1 - f/q_j| \quad (9)$$

and f is the FFT bin center frequency in kHz.

The standard SII calculation assumes that the speech and noise spectra have been measured separately. Let $P(k)$ be the speech power spectrum and $N(k)$ be noise power spectrum computed separately using FFT techniques and where k is the FFT bin index. The SNR for the j th critical band is then given by

$$\text{SNR}(j) = \frac{\sum_{k=0}^K W_j(k) P(k)}{\sum_{k=0}^K W_j(k) N(k)}. \quad (10)$$

In a system that includes distortion, the noise-plus-distortion spectrum is not directly available and must therefore be estimated from the MSC. The speech power spectrum computed using the MSC is given by

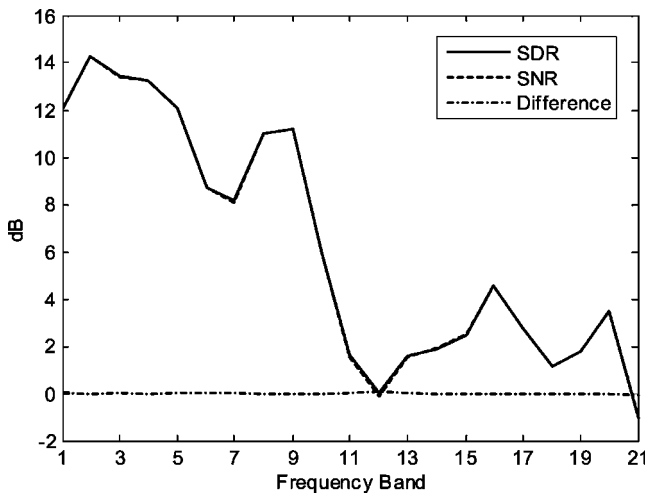


FIG. 2. SDR (solid line), SNR (dashed line), and the difference between them (dot-dashed line) as a function of the frequency band number for the concatenated HINT list 1 sentences. Additive noise low-pass filtered at 900 Hz is present at an SNR of 10 dB.

$$\hat{P}(k) = |\gamma(k)|^2 S_{yy}(k) \quad (11)$$

and the noise power spectrum is given by

$$\hat{N}(k) = [1 - |\gamma(k)|^2] S_{yy}(k), \quad (12)$$

where $S_{yy}(k)$ is the output power spectral density estimated using the FFT. The SDR estimated using the MSC is thus

$$\text{SDR}(j) = \frac{\sum_{k=0}^K W_j(k) |\gamma(k)|^2 S_{yy}(k)}{\sum_{k=0}^K W_j(k) [1 - |\gamma(k)|^2] S_{yy}(k)}. \quad (13)$$

The calculation of Eq. (13) uses the output power spectrum and the coherence, and thus requires access to the input and output signals of the system. However, the system under test can be treated as a “black box;” the procedure does not require knowledge of the distortion mechanism or separate measurements of the noise or distortion signals. The standard SII is computed using the procedures of ANSI S3.5-1997 and the SNR estimate of Eq. (10). The coherence-based SII is computed using the same procedures, but with the speech power spectrum replaced by Eq. (11), the noise power spectrum replaced by Eq. (12), and the SNR replaced by the SDR estimate of Eq. (13).

B. Validation for additive noise

The SDR computed using Eq. (13) and the SNR computed using Eq. (10) should agree for additive noise. That is, the SDR should give the same results as the SNR if only additive noise (and no distortion) is present in the system being measured. If the SDR equals the SNR, the standard and coherence approaches for computing the SII will yield identical results for additive noise.

The SDR and SNR are compared for additive low-pass filtered noise in Fig. 2. The input was a concatenation of the ten sentences in the HINT test list 1 (Nilsson *et al.*, 1994) read by a male talker. The gaps between the sentences were removed. The sentences were digitized at a 44.1-kHz sampling rate and down-sampled to 22.05 kHz. The stimulus duration was approximately 18 s. Figure 2 is for the concat-

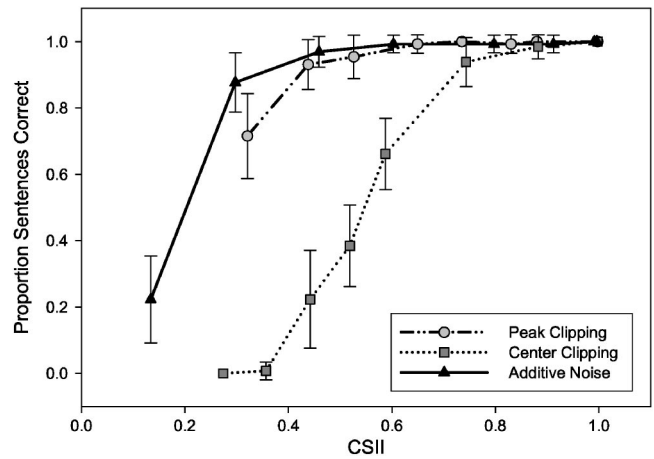


FIG. 3. Proportion of the HINT sentences identified correctly plotted versus the full-level coherence SII calculation for normal-hearing subjects. The triangles, circles, and squares show the average results for the additive noise, peak-clipping, and the center clipping, respectively.

enated sentences with low-pass filtered (one-pole filter at 900 Hz) noise at a 10-dB SNR. The spectra used a FFT size of 32 ms, which is 706 samples at the 22.05-kHz sampling rate. The FFTs were computed using 706-point segments having 50% overlap and weighted with a Hamming window. A new FFT was thus computed every 16 ms, and the FFT had a frequency resolution of 31.23 Hz.

The SNR and SDR are nearly identical, with the largest difference between the two being less than 0.5 dB. Thus the requirement that the SDR and SNR agree for additive noise is satisfied. The SII values computed using the SDR and SNR are identical, being 0.671 for a signal level of 65 dB SPL.

C. Intelligibility predictions

The proportion of the HINT sentences identified correctly is plotted in Fig. 3 as a function of the coherence SII (CSII) for the noise and distortion conditions. The results for the additive speech-shaped noise are in reasonable agreement with the results of Eisenberg *et al.* (1998); an SII of 0.5 corresponds to approximately 90% sentences correct for both sets of data. A CSII of 0.2 corresponds to approximately 35% correct in the results plotted in Fig. 3, while Eisenberg *et al.* (1998) found that an SII of 0.3 gave approximately the same performance.

The intelligibility for peak-clipped speech is similar to that for the additive noise for the same CSII values. The worst case for peak clipping is infinite clipping, which results in a CSII value of approximately 0.3. But even for this worst case there does not appear to be a significant difference between the additive noise and the peak-clipping intelligibility results. For center clipping, however, the intelligibility is significantly worse than for additive noise or peak clipping at the same CSII values. A possible explanation is that center clipping removes a large fraction of the consonant information in the speech signal, but the CSII calculation is dominated by the more-intense portions of the signal and thus primarily reflects modifications of the vowels. Thus the assumption that equal CSII values will give equal intelligibility

is not valid, and the CSII approach cannot accurately predict speech intelligibility for all forms of distortion and noise.

V. THREE-LEVEL CSII

The goal of the new speech intelligibility metric is to derive a single model that accurately predicts the effects of the additive noise, peak clipping, and center clipping used in the experiment. The coherence SII, as implemented above, is inadequate in representing the effects of all three noise and distortion conditions. More information about how both noise and distortion affect the speech signal is needed to predict the subject intelligibility results. This additional information can be provided by dividing the speech signal envelope into three amplitude regions and performing a separate CSII calculation for each region. The three CSII values are then combined to model the intelligibility scores.

A similar multi-level procedure was proposed for estimating perceived speech intelligibility for voice communication systems (CCITT, 1986). In the CCITT (1986) procedure, a synthetic speech signal is divided into overlapping 32-ms windowed segments. The mean-squared level of each segment is computed, and a histogram constructed from the segment magnitudes. Each segment is then assigned to one of four quartiles of the histogram. For the segments in each quartile, the MSC is computed and used to scale the output power spectrum for the quartile to give the speech and noise power spectral estimates. The noise power spectrum is combined with the auditory threshold to give a masking function, the speech power spectrum is converted to sensation levels, and an intermediate score is computed in a manner similar to the SII. The intermediate scores in each quartile are then summed using a weight of 0.556 for the lowest quartile and weights of 0.148 for the upper three quartiles. The weighted sum is then converted to a mean opinion score (MOS) using a nonlinear transformation consisting of the ratio of two logistic functions. Scores predicted using this approach were found to have an average correlation coefficient of 0.97 with subjective ratings of telephone speech processed through various encoding/decoding procedures, additive circuit noise, and quadratic and cubic distortion mechanisms (CCITT, 1987).

A. Coherence SII calculation

To compute the three-level CSII, the speech input signal envelope is divided into three amplitude regions. The calculation uses a block size of 16 ms with a Hamming window and 50% overlap between windowed segments. The magnitude of the signal in each segment is computed and stored over the duration of the sentence. The root-mean squared (RMS) level of the segments is then computed. The high-level segments are those at or above the overall RMS level. The mid-level segments are those between 0 and 10 dB below the overall RMS level, and the low-level segments are those between 10 and 30 dB below the overall RMS level.

The assignment of sentence envelope samples to the different levels is illustrated in Fig. 4 for the first sentence from HINT list 1. Approximately one-third of the segments fall into each of the three level regions. The high-level segments can be considered to be primarily vowels, the mid-level seg-

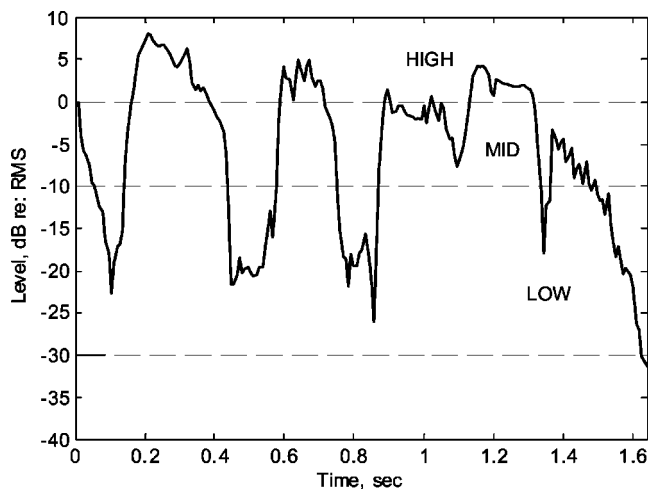


FIG. 4. Envelope for a sentence (HINT list 1, sentence 1) showing the level in dB relative to the RMS level. The dashed lines show the boundaries of the high, mid, and low level regions used for computing the CSII.

ments contain vowel-consonant transitions, and the low-level segments are mostly consonants and pauses. In terms of noise and distortion, the high-level segments will be most strongly affected by peak clipping, while the low-level segments will be most strongly affected by additive noise and center clipping.

For the computation, the output and input signals are first time-aligned. The high-level CSII ($CSII_{High}$) value is then computed by applying the coherence SII calculation procedure, including the auditory threshold, to just the high-level input segments and the corresponding output segments. The high-level MSC is computed using Eq. (6), but the summation is only for those segments classified as high-level input segments. The power spectra used in Eqs. (11)–(13) are similarly restricted to the high-level input segments and the corresponding output segments. The procedure is then repeated to give the mid-level CSII ($CSII_{Mid}$) and the low-level CSII ($CSII_{Low}$) values, after which the high-, mid-, and low-level CSII values are combined to give the prediction of speech intelligibility. In comparison with the CCITT (1986) procedure, the three-level CSII procedure is more computationally efficient since it does not require computing the signal magnitude histogram, and it requires spectrum and coherence calculations over three rather than four signal amplitude regions.

The correlations between the three level CSII values and the overall signal CSII and the subject results are presented in Table I. The correlation between the overall signal CSII and the average intelligibility scores is 0.722. The correlations between the low-level CSII and the scores and between the mid-level CSII and the scores are all higher than the correlations between the overall signal CSII and the scores. Thus the overall signal CSII is not the most useful statistic for describing the effects of the noise and distortion used in the experiment. Furthermore, the correlation between the low-level CSII and the overall signal CSII is only 0.703, indicating that the low-level CSII provides information about the noise and distortion that is absent in the overall signal CSII. Thus the combination of low-, mid-, and high-level

TABLE I. Correlations between the level CSII values, the overall signal CSII, and the average subject intelligibility scores.

	Low-level CSII	Mid-level CSII	High-level CSII	Overall CSII	Average Intelligibility
Low-level CSII	1	0.902	0.508	0.703	0.733
Mid-level CSII		1	0.799	0.922	0.818
High-level CSII			1	0.966	0.626
Overall CSII				1	0.722
Average Intelligibility					1

CSII values provides more information about the noise and distortion and should yield a more-accurate prediction of speech intelligibility and quality than provided by the overall signal CSII.

B. Intelligibility model

The speech intelligibility model uses a linear weighting of the low-, mid-, and high-level CSII values. The weighted sum is then transformed using a logistic (or log-sigmoid) function to give the predicted proportion sentences correct (Gordon-Salant and Fitzgibbons, 1995). The weights are constrained to be positive numbers, and a bias term that adjusts the horizontal position of the curve is also included. The optimum fit of the model to the subject intelligibility scores was found using an unconstrained nonlinear minimization procedure (Nelder and Mead, 1965) to minimize the mean-squared error between the model and the average subject intelligibility scores. Let I_3 denote the intelligibility predicted by the model. The optimum solution is given by

$$c = -3.47 + 1.84CSII_{Low} + 9.99CSII_{Mid} + 0.0CSII_{High},$$

$$I_3 = \frac{1}{1 + e^{-c}}. \quad (14)$$

The intelligibility appears to be dominated by the mid-level CSII, with some input from the low-level CSII. The weight for the high-level CSII is zero, so this term has no apparent effect on intelligibility in the context of this model. The high-level segments are predominantly vowel nuclei, which exhibit little variation over time in comparison with many other speech sounds (Greenberg, 2005), and the contribution of these vowel segments to intelligibility appears to be small (Yoo *et al.*, 2004). The most important term in Eq. (14) is the mid-level CSII, which contains much of the information on envelope transients and spectral transitions that indicate place and manner of articulation (Greenberg, 2005). Preserving the integrity of the speech envelope modulation in the mid-amplitude region appears to be important for both intelligibility (Plomp, 1988; Yoo *et al.*, 2004) and speech quality (Knagenhjelm and Kleijn, 1995). The low-level CSII makes a smaller contribution to intelligibility and, as shown in Table I, much of the information in the low-level region is correlated with that in the mid-level region. The bias value of

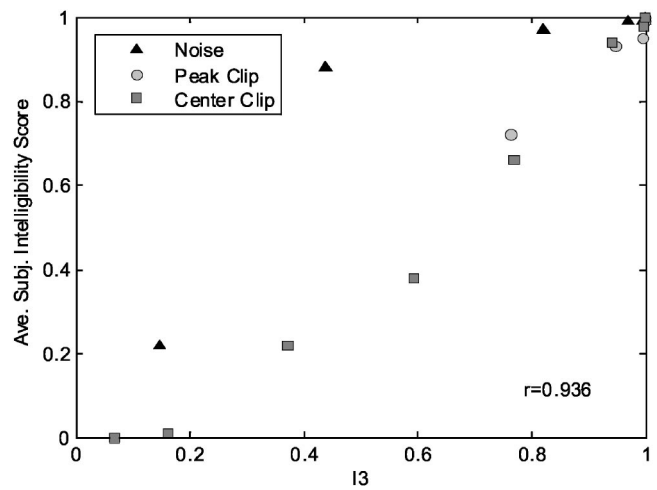


FIG. 5. Proportion of the HINT sentences identified correctly plotted versus the three-level CSII intelligibility predictions I_3 for the normal-hearing subjects. The triangles, circles, and squares show the average results for the additive noise, peak clipping, and the center clipping, respectively.

–3.47 in Eq. (14) provides a lateral shift of the prediction curve that minimizes the error between the prediction and the observed subject intelligibility scores.

The accuracy of the model predictions is shown in Fig. 5. The average intelligibility scores are plotted as a function of I_3 for additive noise, peak clipping, and center clipping. The model tends to underestimate the intelligibility for additive noise and to overestimate the intelligibility for center clipping. The overall accuracy of the model is high, however, with a correlation coefficient between the intelligibility scores and the model predictions of $r=0.94$.

VI. EXPERIMENT II: HEARING-IMPAIRED LISTENERS

The above results indicate that the three-level coherence SII procedure is accurate for normal-hearing listeners. Since an objective of this work is accurate prediction of intelligibility for hearing aids, the procedure must also be shown to work for hearing-impaired listeners. The intelligibility experiment was therefore repeated for a group of hearing-impaired subjects, and the results analyzed using the same three-level coherence SII technique that was used for the normal-hearing subjects.

A. Methods

Nine adult listeners with hearing loss of presumed cochlear origin (mean age=49 years old; age range =23–81 years) participated in experiment 2. (Audiometric thresholds of the listeners with hearing loss are plotted as part of Fig. 7.) Listeners underwent an audiometric evaluation during their initial visit. All of the listeners with hearing loss demonstrated test results that were consistent with cochlear impairment: normal tympanometry, the absence of excessive reflex decay, the absence of air-bone gap exceeding 10 dB at any frequency, and the absence of otoacoustic emissions in regions of threshold loss. All listeners were native English speakers.

As in experiment 1, the hearing-impaired listeners were tested individually in a double-walled sound-treated booth.

The stimuli were the same HINT sentence materials as were used for the normal-hearing subjects. The additive speech-shaped noise, peak-clipping distortion, and center-clipping distortion levels were set using the same procedures as for the normal-hearing listeners. The corrupted sentences were then equalized in level so that they had the same long-term 65-dB SPL amplitude as the unprocessed stimuli. The corrupted and level-adjusted sentences were then amplified using the NAL-R gain rule (Byrne and Dillon, 1986) computed for each hearing loss at the audiometric frequencies from 250 to 4000 Hz. The gains in dB were interpolated across frequency and used to design a 128-point linear-phase finite impulse response (FIR) filter through which the stimuli were processed. The sentences were presented monaurally to the listener's test ear. The equipment and headphones were the same as for the normal-hearing subjects.

B. Intelligibility scores

The proportion of the sentences identified correctly for the additive noise, peak-clipping distortion, and center-clipping distortion is plotted in Fig. 6 for each of the nine hearing-impaired subjects. The general trend of the data is poorer performance as the amount of noise or distortion is increased, with the spread in the data also increasing as the noise or distortion increases. The speech recognition score for all of the subjects is 80% or better for no noise or distortion, and drops to close to zero for the worst-case noise and center clipping conditions. The overall trends in speech-recognition performance for the hearing-impaired listeners agree with those in the normal-hearing listener data presented in Fig. 1.

In general, the subjects with the greatest hearing loss have the poorest performance. For example, subject V (solid triangles) has a flat loss ranging from 30 dB at 250 Hz to an average of 50 dB above 2000 Hz and has nearly the best performance of the hearing-impaired subjects. Conversely, subject O (open circles) has a steeply sloping loss with over 70 dB of loss above 2000 Hz, and has uniformly poor performance.

C. Hearing-impaired CSII

The predicted intelligibility for the hearing-impaired subjects is given by the same I_3 calculation procedure as used for the normal-hearing subjects. The speech segments are each assigned to one of three level regions, and the coherence computed separately for the groups of low-, mid-, and high-level speech segments. The three coherence values are then combined using the offset and weights given in Eq. (14), followed by the logistic function transformation. The numerical coefficients used in Eq. (14) for the hearing-impaired listeners were exactly the same as those used for the normal-hearing listeners, and the upward spread of masking included in the SII calculation was as specified in the standard procedure for normal ears. The level distortion factor in the SII calculation was not used.

The only difference between the calculations for the normal-hearing and hearing-impaired subjects is that the calculation for the hearing-impaired listeners includes the hear-

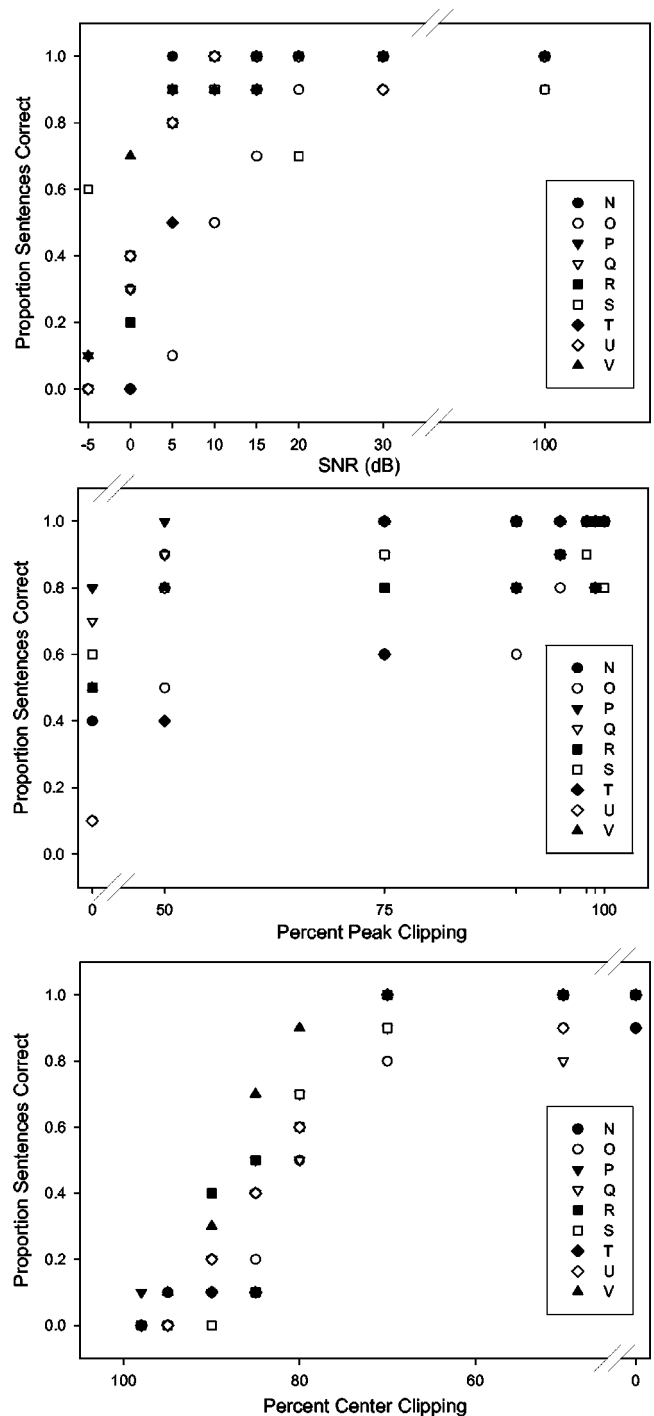


FIG. 6. Proportion of the HINT sentences identified correctly plotted for each hearing-impaired subject for the additive speech-shaped noise, peak clipping, and center clipping distortion.

ing loss and the NAL-R amplification. These two adjustments affect the SDR as a function of frequency. The audiogram in the SII calculations is represented as an equivalent internal noise source, so the hearing loss reduces the SDR. The NAL-R amplification increases the signal gain in the regions of hearing loss and thus partly restores the SDR. The amplification, however, provides imperfect compensation; the SDR for all three signal levels is therefore reduced by the hearing loss, with the greatest effect occurring for the low-level CSII. The predicted speech intelligibility is there-

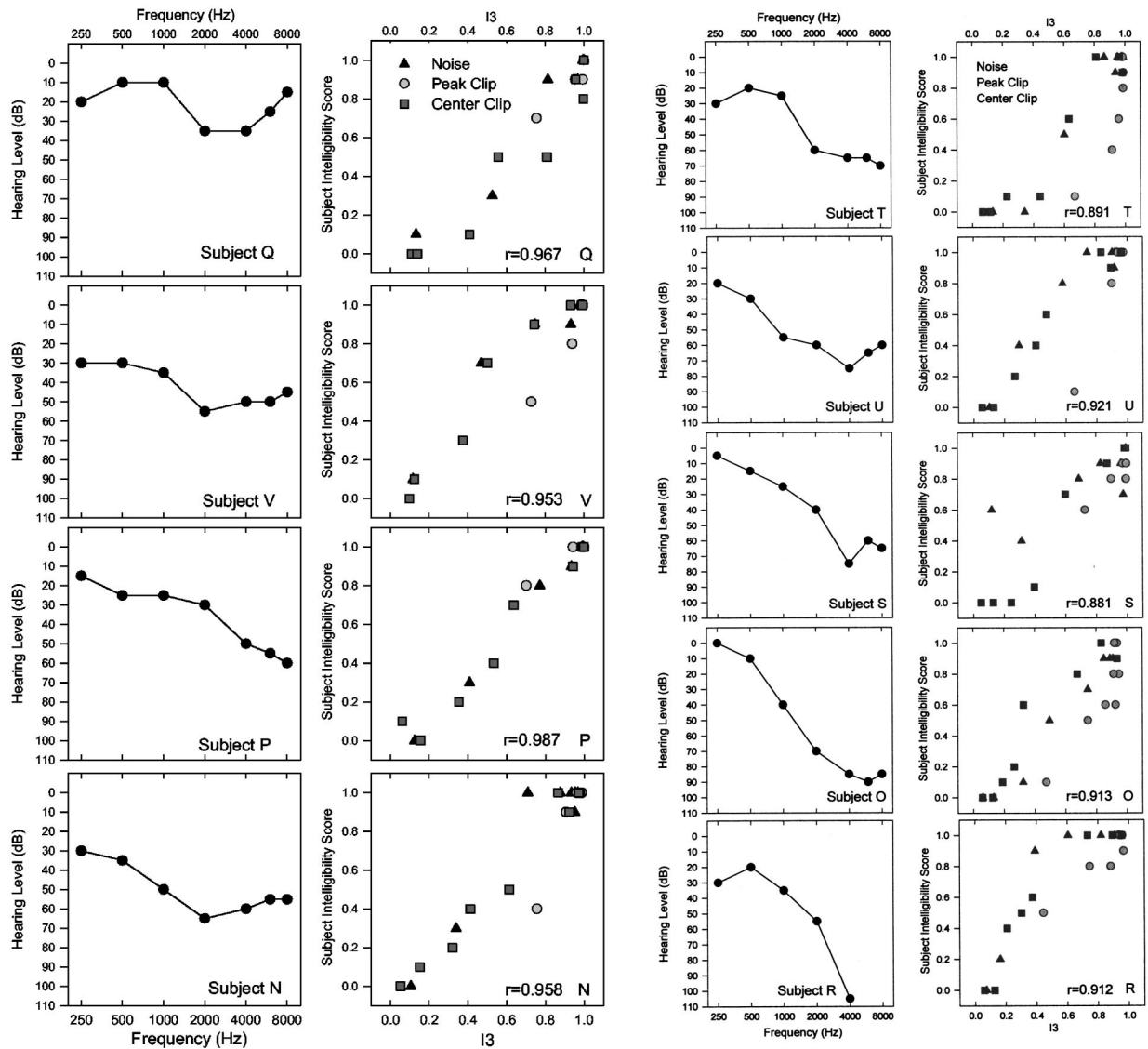


FIG. 7. Audiograms and the three-level CSII intelligibility predictions I_3 for each hearing-impaired subject. The triangles, circles, and squares show the average results for the additive noise, peak-clipping, and the center clipping, respectively.

fore reduced both by the presence of noise and distortion in the signal and by the reduced audibility of speech sounds caused by the hearing loss.

The individual subject results are shown in Fig. 7, where the audiogram is presented for each subject along with the plot of the observed versus predicted intelligibility. The subjects are ordered by increasing severity of the high-frequency hearing loss. As explained above, the predictions for each hearing-impaired subject use the coefficients derived for the normal-hearing subjects in conjunction with the individual audiogram and NAL-R gain. The correlation coefficients between the predicted and observed intelligibility scores range from 0.88 (subject S) to 0.99 (subject P). There appears to be a slight trend of decreasing prediction accuracy with increasing hearing loss, but the I_3 predictions are quite accurate even for the most severe losses included in this study.

In Fig. 8, the average hearing-impaired intelligibility scores are plotted against the average I_3 predictions for each noise or distortion condition. The correlation coefficient between the average predictions and subject scores is 0.98. The

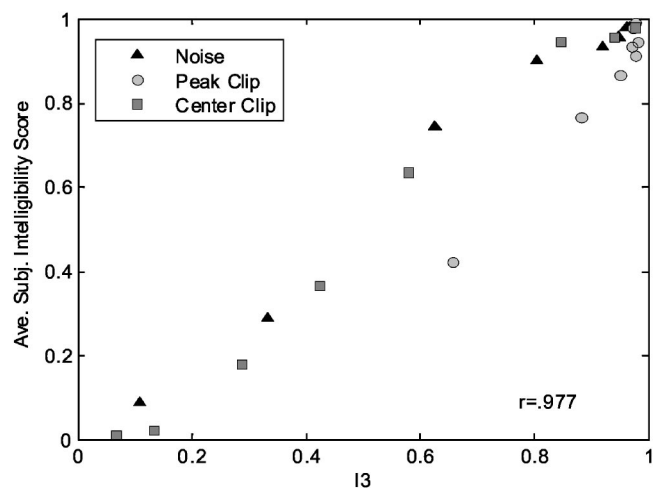


FIG. 8. Average proportion of the HINT sentences identified correctly plotted versus the three-level CSII intelligibility predictions I_3 for the hearing-impaired subjects. The triangles, circles, and squares show the average results for the additive noise, peak-clipping, and the center clipping, respectively.

accuracy of the intelligibility predictions is high for all three forms of noise and distortion used in this experiment despite the widely different distortion mechanisms. Peak clipping predominantly affects the high-level speech sounds, with reduced effects on the mid-level sounds, and noise and center clipping primarily affect the low-level speech sounds, again with reduced effects on the mid-level sounds. Because the I_3 calculation procedure divides the speech into high-, mid-, and low-level regions, it can effectively deal with the reduced audibility of speech information occurring at different speech amplitude levels for the noise and distortion used in this study.

VII. DISCUSSION

The experiments described in this paper deal with two issues. The first issue is extending the SII approach to include distortion as well as noise, and the second is predicting speech intelligibility for hearing-impaired listeners.

A. Distortion

Replacing the SNR in the SII calculation with the SDR computed from the coherence is an intuitively obvious extension of the principles underlying the SII. Authors of papers using coherence to measure distortion in hearing aids (Preves *et al.*, 1989; Dyrland, 1989; Preves, 1990; Kates, 1992; Fortune and Preves, 1992; Schneider and Jamieson, 1995; Kates, 2000) have made the assumption that there is a direct relationship between a reduction in coherence and a reduction in intelligibility. For any one type of noise or distortion, there will be a monotonic relationship between the increase in the amount of distortion and the reduction in the SDR. Thus earlier experiments that looked at just one type of distortion, such as peak clipping (Steeneken and Houtgast, 1980; Crain and van Tasell, 1994; Kates and Kozma-Spytek, 1994), can produce an accurate mapping of the SDR to intelligibility or sound quality for the distortion condition considered in the study. But as shown in this paper, those mappings may not be accurate if more than one type of distortion is considered. It is much more difficult to produce a relationship that is accurate for both center clipping, which most strongly affects the lower signal amplitude regions, and peak clipping, which most strongly affects the higher signal amplitude regions. The three-level coherence SII, on the other hand, combines the effects of the three different noise and distortion mechanisms considered in this paper into a single relationship that is accurate for all of them.

The three-level CSII approach divides the speech into low-, mid-, and high-amplitude regions. The three level CSII calculations provide much more information about the effects of the distortion on the speech than can be obtained by the SNR or the full-level CSII alone. In particular, the high-level CSII, which dominates the full-level CSII, has a weight of zero in the three-level calculation and thus no observed effect on intelligibility even though the high-level CSII is greatly reduced by peak clipping. Most of the speech information appears to be conveyed within a narrow 10-dB dynamic range from the RMS level of the sentence to 10 dB below, with a smaller contribution from the lower speech levels. The three-level CSII approach thus provides enough

information to accurately model the reduction in intelligibility resulting from all three types of noise and distortion considered in this study, and does so in a manner consistent with observations on speech production and perception (Greenberg, 2005; Yoo *et al.*, 2004). Increasing the number of levels to more than three may improve the accuracy of the intelligibility predictions, but it would also increase the number of degrees of freedom in the model and would thus increase the chances of over-fitting (fitting individual data points rather than the overall structure of the data) the limited amount of data available from the experiments.

B. Hearing-impaired listeners

Many researchers have tried to predict speech intelligibility in hearing-impaired listeners. The approaches typically taken in modeling the effects of hearing loss are the addition of a desensitization factor to the SII calculation (Pavlovic *et al.*, 1986; Ching *et al.*, 1998; Magnusson *et al.*, 2001) or the inclusion of increased spread of masking (Ludvigsen, 1987; Rankovic, 1998) or individually measured masking functions (Hornsby and Ricketts, 2003). The desensitization factor is equivalent to a reduction in the SNR used in the SII calculation, with the amount of the reduction based on empirical observations of speech intelligibility in hearing-impaired listeners compared with those having normal hearing. Increasing the masking used in the SII calculation, either through increased spread of masking or the use of individual masking functions, has a similar effect in that it reduces the effective SNR.

The three-level CSII gives an alternative approach to modeling the effects of hearing loss. The CSII calculation for the hearing-impaired subjects requires no modification to the calculation procedure or masking functions; the audiogram and amplification used for the impaired ear are the only necessary inputs, after which the calculation for the hearing-impaired listeners proceeds exactly as for the normal-hearing listeners. The difference is that the three-level CSII incorporates a much more complete model of the effects of hearing loss on the audibility of speech. The low-level CSII is strongly affected by the hearing loss; the low-level speech sounds are much closer to threshold in the impaired than in the normal ear, and the low-level CSII is therefore much lower in the impaired ear even without noise or distortion. The addition of noise or distortion will reduce the low-level CSII even further, thus quickly removing the speech information provided in the low-level signal region. The mid-level CSII will experience similar effects, but with reduced severity given the higher signal level. The high-level CSII is affected the least, but the high-level CSII is not a factor in the I_3 calculation of Eq. (14).

These results indicate that audibility is an important factor in the difference in speech intelligibility between normal-hearing and hearing-impaired listeners. No modifications were made to the auditory filter shapes, upward spread of masking, or the importance functions used in computing I_3 for the hearing-impaired subjects, yet very accurate predictions of intelligibility were obtained over a wide range of hearing losses by incorporating just the audiogram. The SNR used in the standard SII calculation reflects the level of the

noise relative to the speech peaks and gives an imprecise indication of the effects of the noise on the low-level speech components. The desensitization factor or increased masking proposed previously for hearing-impaired listeners can thus be seen as attempts to take an imperfect measure of speech audibility, the full-level SNR, and modify it to reflect the greater loss of audibility of low-level speech information experienced by the impaired ear.

The impaired auditory system exhibits processing deficits in addition to the elevated auditory threshold. These deficits include recruitment (Reger, 1936; Allen *et al.*, 1990), broadened auditory filters (Tyler, 1986; Dubno and Schaefer, 1991), increased spread of masking (Gagné, 1983; Glasberg and Moore, 1986), and poorer temporal resolution (Oxenham and Bacon, 2003). It is possible that some of these effects are subsumed into the three-level coherence SII approach. For example, the loss of information in the low-level signal segments may be highly correlated with the audiogram, and thus predicted by the I_3 calculation, but could still be caused by poor temporal processing or other deficits in the impaired ear. The predictive value of the SII, or of the three-level coherence SII extension, is based on correlations and does not necessarily indicate a causal relationship. It is also possible that certain types of interference that were not considered in this paper, such as temporally fluctuating noise, may cause reductions in intelligibility for hearing-impaired listeners (Festen and Plomp, 1990; Peters *et al.*, 1998) that are not predicted by the three-level coherence SII model.

C. Limitations and extensions

One limitation of the work reported in this paper is that the noise was broadband, and the distortion mechanisms were also applied to the broadband signals. These test conditions are appropriate for many communication systems. In hearing aids, however, frequency-dependent amplification is used, which can result in different amounts of noise or distortion in different frequency regions. Previous results (Kates and Kozma-Spytek, 1994; Kozma-Spytek *et al.*, 1996; Tan *et al.*, 2004) have shown that approaches based on the principle of coherence are accurate in predicting speech quality for frequency-dependent distortion conditions as well as for broadband distortion. However, experimental verification is needed to show that the three-level coherence SII approach is also valid for frequency-dependent distortion. Experiments are also needed to extend the range of noise and distortion conditions considered. Quantization noise can be caused by the digital processing in a hearing aid or communication system, and transducer problems such as displacement limiting and nonuniform driver magnetic fields can also cause distortion. An additional issue is the accuracy of the three-level coherence SII approach for fluctuating noise. Additional experiments are needed to validate and refine the metric for the wide range of noise and distortion conditions that occur in real life.

The three-level CSII approach should also lead to an improved procedure for predicting the quality of distorted speech. There is a strong correlation between intelligibility and sound quality (Preminger and Van Tasell, 1995; Eisenberg *et al.*, 1998), which would lead one to expect a similarly

high correlation between the intelligibility predicted by I_3 and sound quality for comparable noise and distortion conditions. While preliminary results are encouraging (Kates and Arehart, 2004), additional experiments are needed to extend the three-level coherence SII approach to produce an accurate speech quality metric for normal-hearing and hearing-impaired listeners.

VIII. CONCLUSIONS

Speech intelligibility is the primary concern in hearing aids and many communication systems. The SII has been developed and validated for additive noise and for bandwidth reduction, but the output of hearing aids is affected by distortion as well as by noise. The coherence between the input and the output of the device under test provides a tool for calculating the combined effects of noise and distortion in the processed output. It is therefore attractive to modify the SII calculation procedure to use the SDR computed from the coherence instead of the SNR, and thus include distortion along with noise.

Intelligibility was determined for broadband additive speech-shaped noise, peak clipping, and center clipping for normal-hearing and hearing-impaired subjects. A simple modification of the SII, using the SDR in place of the SNR but leaving the remainder of the procedure unchanged, was shown to be inadequate in predicting the intelligibility data for the normal-hearing listeners. The estimated intelligibility for peak clipping was similar to that for noise, but center clipping had much lower intelligibility for the same coherence SII values.

Much better accuracy was obtained by dividing the speech signal segments into three level regions. The high-level region consists of those segments at or above the sentence RMS level. The mid-level region ranges from the RMS sentence level to 10 dB below, and the low-level region ranges from RMS - 10 dB to RMS - 30 dB. The coherence SII is then computed separately for the segments in each level region. A linear combination of the CSII values, followed by a logistic function transformation, models the normal-hearing subject intelligibility scores with a correlation coefficient of 0.94.

The intelligibility experiment was then repeated for a group of hearing-impaired listeners. The distorted speech was amplified using the NAL-R gains computed for each hearing loss, and the three-level CSII calculation was adjusted for the hearing loss by including the audiogram and the NAL-R gain in the constituent CSII calculations. The intelligibility predictions were accurate across the entire range of hearing losses used in the experiment, and the average across subjects for the different noise and distortion conditions had a correlation coefficient of 0.98. The three-level CSII thus appears to be an accurate procedure for estimating intelligibility for the conditions of distortion and noise considered in this paper, and it is as effective for hearing-impaired listeners as it is for those with normal hearing.

ACKNOWLEDGMENTS

Author KHA was supported by the Whitaker Foundation. The information provided in this paper was also supported in part by Grant/Cooperative Agreement Number UR3/CCU824219 from the Centers for Disease Control and Prevention (CDC). The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of CDC. The authors thank Rosalinda Baca, Denise Kidd, and Jessica Rossi-Katz for assistance in the data collection, and Amit Das for development of the software used for the listener tests. The authors also thank Professor Brian Moore for helpful comments on earlier drafts of this paper.

- Allen, J. B., Hall, J. L., and Jeng, P. S. (1990). "Loudness growth in $\frac{1}{2}$ -octave bands—a procedure for the assessment of loudness," *J. Acoust. Soc. Am.* **88**, 745–753.
- ANSI S3.22-1996 (1996). *American National Standard: Specification of Hearing Aid Characteristics* (American National Standards Institute, New York).
- ANSI S3.42-1992 (1992). *American National Standard: Testing Hearing Aids with a Broadband Noise Signal* (American National Standards Institute, New York).
- ANSI S3.5-1997 (1997). *American National Standard: Methods for the Calculation of the Speech Intelligibility Index* (American National Standards Institute, New York).
- ANSI S3.6-1989 (1989). *American National Standard: Specification for Audiometers* (American National Standards Institute, New York).
- Byrne, D., and Dillon, H. (1986). "The National Acoustics Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid," *Ear Hear.* **7**, 257–265.
- Carter, G. C., Knapp, C. H., and Nuttall, A. H. (1973). "Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing," *IEEE Trans. Audio Electroacoust.* **21**, 337–344.
- CCITT (1986). "Objective evaluation of non-linear distortion effects on voice transmission quality," CCITT Study Group XII, Communication XII No. 8.
- CCITT (1987). "Re-evaluation of the objective method for measurement of non-linear distortion," CCITT Study Group XII, Communication XII-175-E.
- Ching, T. Y. C., Dillon, H., and Byrne, D. (1998). "Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification," *J. Acoust. Soc. Am.* **103**, 1128–1140.
- Crain, T., and Van Tasel, D. J. (1994). "Effect of peak clipping on speech recognition threshold," *Ear Hear.* **15**, 443–453.
- Dubno, J. R., and Schaefer, A. B. (1991). "Frequency selectivity for hearing-impaired and broadband-noise masked normal listeners," *Q. J. Exp. Psychol.* **43**, 543–564.
- Dyrlund, O. (1989). "Characterization of non-linear distortion in hearing aids using coherence: A pilot study," *Scand. Audiol.* **18**, 143–148.
- Eisenberg, L. S., Dirks, D. D., Takayanagi, S., and Martinez, A. S. (1998). "Subjective judgments of clarity and intelligibility for filtered stimuli with equivalent speech intelligibility index predictions," *J. Speech Lang. Hear. Res.* **41**, 327–339.
- Festen, J. M., and Plomp, R. (1990). "Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing," *J. Acoust. Soc. Am.* **88**, 1725–1736.
- Fletcher, H., and Galt, R. H. (1950). "The perception of speech and its relation to telephony," *J. Acoust. Soc. Am.* **22**, 89–151.
- Fortune, T. W., and Preves, D. A. (1992). "Hearing aid saturation and aided loudness discomfort," *J. Speech Hear. Res.* **35**, 175–185.
- French, N. R., and Steinberg, J. C. (1947). "Factors governing the intelligibility of speech sounds," *J. Acoust. Soc. Am.* **19**, 90–119.
- Gagné, J.-P. (1983). "Excess masking among listeners with a sensorineural hearing loss," *J. Acoust. Soc. Am.* **83**, 2311–2321.
- Glasberg, B. R., and Moore, B. C. J. (1986). "Auditory filter shapes with unilateral and bilateral cochlear impairments," *J. Acoust. Soc. Am.* **79**, 1020–1033.
- Gordon-Salant, S., and Fitzgibbons, P. (1995). "Comparing recognition of distorted speech using an equivalent signal-to-noise ratio," *J. Speech Hear. Res.* **38**, 706–713.
- Greenberg, S. (2005). "A multi-tier framework for understanding spoken language," to appear in *Listening to Speech: An Auditory Perspective*, edited by S. Greenberg and W. Ainsworth (Erlbaum, Mahwah, NJ).
- Hornsby, B. W. Y., and Ricketts, T. A. (2003). "The effects of hearing loss on the contributions of high- and low-frequency speech information to speech understanding," *J. Acoust. Soc. Am.* **113**, 1706–1717.
- Houtgast, T., and Steeneken, H. J. M. (1985). "A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria," *J. Acoust. Soc. Am.* **77**, 1069–1077.
- Kates, J. M. (1987). "The short-time articulation index," *J. Rehabil. Res. Dev.* **24**, 271–276.
- Kates, J. M. (1992). "On using coherence to measure distortion in hearing aids," *J. Acoust. Soc. Am.* **91**, 2236–2244.
- Kates, J. M. (2000). "Cross-correlation procedures for measuring noise and distortion in hearing aids," *J. Acoust. Soc. Am.* **107**, 3407–3414.
- Kates, J. M., and Arehart, K. H. (2004). "A metric for evaluating speech intelligibility and quality in hearing aids," *J. Acoust. Soc. Am.* **116**, 2536–2537.
- Kates, J. M., and Kozma-Spytek, L. (1994). "Quality ratings for frequency-shaped peak-clipped speech," *J. Acoust. Soc. Am.* **95**, 3586–3594.
- Knagenhjelm, H. P., and Kleijn, W. B. (1995). "Spectral dynamics is more important than spectral distortion," *Proc. IEEE Int. Conf. on Acoust. Speech and Sig. Proc.*, Detroit.
- Kozma-Spytek, L., Kates, J. M., and Revoile, S. G. (1996). "Quality ratings for frequency-shaped peak-clipped speech: Results for listeners with hearing loss," *J. Speech Hear. Res.* **39**, 1115–1123.
- Kryter, K. D. (1962). "Methods for the calculation and use of the Articulation Index," *J. Acoust. Soc. Am.* **34**, 1689–1697.
- Licklider, J. C. R. (1946). "Effects of amplitude distortion upon the intelligibility of speech," *J. Acoust. Soc. Am.* **18**, 429–434.
- Ludvigsen, C. (1987). "Prediction of speech intelligibility for normal-hearing and cochlearly hearing-impaired listeners," *J. Acoust. Soc. Am.* **82**, 1162–1171.
- Magnusson, L., Karlsson, M., and Leijon, A. (2001). "Predicted and measured speech recognition performance in noise with linear amplification," *Ear Hear.* **22**, 46–57.
- Moore, B. C. J., and Glasberg, B. R. (1983). "Suggested formulas for calculating auditory-filter bandwidths and excitation patterns," *J. Acoust. Soc. Am.* **74**, 750–753.
- Nelder, J. A., and Mead, R. (1965). "A simplex method for function minimization," *Comput. J.* **7**, 308–313.
- Nilsson, M., Soli, S. D., and Sullivan, J. (1994). "Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise," *J. Acoust. Soc. Am.* **95**, 1085–99.
- Oxenham, A. J., and Bacon, S. P. (2003). "Cochlear compression: Perceptual measures and implications for normal and impaired hearing," *Ear Hear.* **24**, 352–366.
- Palmer, C. V., Killion, M. C., Wilber, L. A., and Ballard, W. J. (1995). "Comparison of two hearing aid receiver-amplifier combinations using sound quality judgments," *Ear Hear.* **16**, 587–598.
- Pavlovic, C. V. (1984). "Use of the articulation index for assessing residual auditory functions in listeners with sensorineural hearing impairment," *J. Acoust. Soc. Am.* **75**, 1253–1258.
- Pavlovic, C. V., Studebaker, G. A., and Sherbecoe, R. L. (1986). "An articulation index based procedure for predicting the speech recognition performance of hearing-impaired individuals," *J. Acoust. Soc. Am.* **80**, 50–57.
- Payton, K. L., and Braid, L. D. (1999). "A method to determine the speech transmission index from speech waveforms," *J. Acoust. Soc. Am.* **106**, 3637–3648.
- Peters, R. W., Moore, B. C. J., and Baer, T. (1998). "Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people," *J. Acoust. Soc. Am.* **103**, 577–587.
- Plomp, R. (1988). "The negative effect of amplitude compression in multi-channel hearing aids in the light of the modulation transfer function," *J. Acoust. Soc. Am.* **83**, 2322–2327.
- Preminger, J. E., and Van Tasell, D. J. (1995). "Quantifying the relation between speech quality and speech intelligibility," *J. Speech Hear. Res.* **38**, 714–725.
- Preves, D. A. (1990). "Expressing hearing aid noise and distortion with coherence measurements," *Am. Speech Hear. Assoc.* **32**, 56–59.

- Preves, D. A., Beck, L. B., Burnett, E. D., and Teder, H. (1989). "Input stimuli for obtaining frequency responses of automatic gain control hearing aids," *J. Speech Hear. Res.* **32**, 189–194.
- Rankovic, C. M. (1998). "Factors governing speech reception benefits of adaptive linear filtering for listeners with sensorineural hearing loss," *J. Acoust. Soc. Am.* **103**, 1043–1057.
- Reger, S. N. (1936). "Differences in loudness response of the normal and hard-of-hearing at intensity levels slightly above threshold," *Ann. Otol. Rhinol. Laryngol.* **45**, 1029–1036.
- Rhebergen, K. S., and Versfeld, N. J. (2004). "An SII-based approach to predict the speech intelligibility in fluctuating noise for normal-hearing listeners," *J. Acoust. Soc. Am.* **115**, 2394.
- Schneider, T., and Jamieson, D. G. (1995). "Using maximum length sequence coherence for broadband distortion measurements in hearing aids," *J. Acoust. Soc. Am.* **97**, 2282–2292.
- Steeneken, H. J. M., and Houtgast, T. (1980). "A physical method for measuring speech-transmission quality," *J. Acoust. Soc. Am.* **67**, 318–326.
- Tan, C.-T., Moore, B. C. J., Zacharov, N., and Matilla, V.-M. (2004). "Predicting the perceived quality of nonlinearly distorted music and speech signals," *J. Audio Eng. Soc.* **52**, 699–711.
- Tyler, R. S. (1986). "Frequency resolution in hearing-impaired listeners," *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London), pp. 309–371.
- Yoo, S., Boston, R., Durrant, J. D., Kovacyk, K., Karn, S., Shaiman, S., El-Jaroudi, A., and Li, C.-C. (2004). "Relative energy and intelligibility of transient speech components," *Proc. EUSIPCO, Vienna*, 6–10 September.
- Zwicker, E. (1961). "Subdivision of the audible frequency range into critical bands," *J. Acoust. Soc. Am.* **33**, 248.
- Zwicker, E., and Terhardt, E. (1980). "Analytical expressions for critical-band rate and critical bandwidth as a function of frequency," *J. Acoust. Soc. Am.* **68**, 1523–1525.

Statistical properties of infant-directed versus adult-directed speech: Insights from speech recognition^{a)}

Katrin Kirchhoff^{b)} and Steven Schimmel

Department of Electrical Engineering, Box 352500, University of Washington, Seattle, Washington 98195

(Received 18 April 2004; revised 20 January 2005; accepted 21 January 2005)

Previous studies have shown that infant-directed speech ('motherese') exhibits overemphasized acoustic properties which may facilitate the acquisition of phonetic categories by infant learners. It has been suggested that the use of infant-directed data for training automatic speech recognition systems might also enhance the automatic learning and discrimination of phonetic categories. This study investigates the properties of infant-directed vs. adult-directed speech from the point of view of the statistical pattern recognition paradigm underlying automatic speech recognition. Isolated-word speech recognizers were trained on adult-directed vs. infant-directed data sets and were tested on both matched and mismatched data. Results show that recognizers trained on infant-directed speech did not always exhibit better recognition performance; however, their relative loss in performance on mismatched data was significantly less severe than that of recognizers trained on adult-directed speech and presented with infant-directed test data. An analysis of the statistical distributions of a subset of phonetic classes in both data sets showed that this pattern is caused by larger class overlaps in infant-directed speech. This finding has implications for both automatic speech recognition and theories of infant speech perception. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1869172]

PACS numbers: 43.72.Ne, 43.71.-k, 43.71.Ft [DOS]

Pages: 2238–2246

I. INTRODUCTION

Many studies on infant speech perception have analyzed the acoustic-phonetic properties of infant-directed speech (also called "motherese" or "parentese"). It has been observed that infant-directed speech is generally slower in tempo and exhibits increased segmental duration, longer pauses, and more pronounced pitch contours (Grieser and Kuhl, 1988; Fernald *et al.*, 1989; Kuhl *et al.*, 1997). Most importantly, it has been shown that motherese is characterized by overemphasized acoustic-phonetic contrasts: compared to adult-directed (AD) speech, vowels in infant-directed (ID) speech have mean formant values closer to the outermost points of the vowel triangle and are thus acoustically more distinct (Andruski and Kuhl, 1996; Kuhl *et al.*, 1997; Burnham *et al.*, 2002). Infants have been shown to prefer infant-directed over adult-directed speech (Cooper and Aslin, 1994; Fernald and Kuhl, 1987), and a recent study (Liu *et al.*, 2003) has found a correlation (though not necessarily a causal link) between the clarity of mothers' speech as measured by vowel space expansion and infants' phonetic discrimination abilities. It has also been demonstrated that auditory input with overemphasized acoustic-phonetic contrasts is helpful in second-language learning (Protopapas and Calhoun, 2000; Hazan and Simpson, 2000) and that it is more intelligible to individuals with hearing impairments (Picheny *et al.*, 1986; Payton *et al.*, 1994). For these reasons, it has often been suggested that ID-style speech might also be useful for training automatic speech recognition (ASR) systems. It is generally recognized that the amount of train-

ing data is of primary importance in developing good ASR systems; however, a recent study (Moore, 2003) which extrapolates state-of-the-art ASR results to larger training sets suggests that ASR performance would still be far from human performance even if training sets were increased by orders of magnitude. An alternative to simply using more data might be to use "better" data, e.g., more clearly articulated acoustic data. The present study explores this hypothesis by training an isolated word recognizer on adult-directed and infant-directed data sets, respectively, and analyzing its performance under identical versus mismatched test conditions.

In addition to assessing the benefit of ID-style speech for ASR we are interested in exploring the ASR framework as a computational modeling technique for infant speech perception, in particular phonetic category learning. The development of a computational model whose predictions match data gathered from human speech perception experiments could prove valuable in predicting the outcome of future experiments through simulations. A good computational model allows one to observe the effect of adjusting individual experimental variables without having to conduct expensive perceptual studies. The framework of ASR can potentially make a significant contribution towards this goal. First, ASR relies on a suite of statistical modeling techniques that have been fine-tuned to speech (e.g., perceptually inspired signal processing and temporal modeling algorithms) and are more advanced than many of the computational models of speech perception proposed in the past. Second, the ASR community has developed automated methods for collecting, preprocessing and "cleaning" speech data (e.g., high-accuracy automatic segmentation) which enable speech

^{a)}Portions of this work were presented at the 146th Meeting of the Acoustical Society of America, Austin, Texas, November 2003.

^{b)}Electronic mail: katrin@ee.washington.edu

researchers to analyze large speech data sets more rapidly and efficiently.

Previous computational or numerical models of speech perception have mainly focused on predicting experimental data using techniques such as logistic regression and discriminant analysis (e.g., Hillenbrand *et al.*, 1995; Nearey, 1997) or connectionist models (Protopapas, 1999; Damper and Hanard, 2000; Guenther and Bohland, 2002). Recently, the idea of using ASR techniques for modeling human speech perception and language acquisition has attracted increased attention. Scharenborg *et al.*, (2002, 2003), for instance, have investigated the problem of word segmentation in human speech perception using an ASR decoder as a modeling tool. They showed, in a small word discrimination experiment, that the segmentations selected by an ASR decoder matched the outcome of human word segmentation experiments. In de Boer and Kuhl (2003) the learnability of infant-directed and adult-directed speech datasets was investigated using a computer model. A mixture of Gaussians was fitted to samples of the vowels /i:/, /u:/ and /ɒ/ obtained from ID and AD speech, respectively. The resulting means of the learned categories were then compared to the expected means for those vowels. Since the means learned from ID speech matched the reference values more closely, it was concluded that the ID speech has better learnability. However, this study did not evaluate the generalizability of the models by applying them to a separate test set. A fundamental concept of statistical learning, however, is the capability of the trained models to classify unseen samples that are not present in the training set.

The present study investigates the benefit of ID versus AD training data for the formation of phonetic category models which are subsequently applied to unseen instances of those categories. The ability of the trained models to generalize to novel data is measured in terms of their classification error rate. Unlike most previous work, which has focused on the analysis of individual phonetic categories (e.g., vowels), this study also investigates the recognition of entire words.

The remainder of this paper is structured as follows: in Sec. II we give a more detailed explanation of the rationale for this work. The data is described in Sec. III, and experiments and results are presented in Sec. IV. Section V concludes.

II. RATIONALE

In order to assess the potential benefit of ID-style speech for training ASR systems, it is necessary to take a closer look at its properties and view them in light of the standard phonetic classification procedures employed in state-of-the-art ASR systems.

As mentioned above, ID speech is characterized by a greater distance between vowel class means measured in formant space. In addition, the variances associated with individual classes are often greater than in adult-directed speech (Kuhl *et al.*, 1997; de Boer and Kuhl, 2003). These properties have been observed in a number of studies of vowel patterns in ID speech, in American English as well as other languages (Chinese, Swedish and Russian) (Kuhl *et al.*,

1997; Liu *et al.*, 2003). Phonetic overspecification has also been observed for stop consonants in the form of longer voice onset times in ID speech addressed to infants aged 11–14 months (Sundberg, 2001). Most of these studies have focused on monosyllabic content words with stressed vowels. In a study of both content and function words in ID speech, van de Weijer (2001) confirmed the observation of an enlarged vowel space for content words but found the opposite pattern in function words.

Studies of infant speech perception, in particular word segmentation, have shown that infants respond to statistical regularities in the speech input (e.g., Jusczyk *et al.*, 1994; Saffran *et al.*, 1996). The assumption that some form of statistical learning underlies the acquisition of phonetic categories as well (Holt *et al.*, 1998) has recently been demonstrated empirically: Maye *et al.* (2002) showed that infants exposed to speech samples from a bimodal distribution of stops (with the two peaks representing voicedness and voicelessness, respectively) were capable of discriminating new samples from the endpoints of those distributions, whereas infants exposed to a unimodal training distribution were not. The precise nature of such a learning mechanism remains as yet obscure, but it is apparent that spectral discrimination as well as temporal information integration must be involved. However, computational models of infant speech perception that explicitly incorporate temporal information processing are rare. In order to better assess the potential contributions of the ASR framework to this field, we briefly review the statistical models used in present-day ASR systems.

A. Acoustic modeling in automatic speech recognition

Current ASR systems are based on a statistical pattern recognition framework. Recognizers attempt to find the best word sequence W^* given a sequence of acoustic observations O :

$$W^* = \arg \max_W P(W|O) \quad (1)$$

The probability $P(W|O)$ is computed using Bayes' rule:

$$P(W|O) = \frac{P(O|W)P(W)}{P(O)} \propto P(O|W)P(W). \quad (2)$$

The *language model* $P(W)$ gives the prior probability of a word sequence W . Since our focus is on acoustic classification and isolated word recognition, properties of the language model are of no further concern in this study. The *acoustic model* $P(O|W)$ maps acoustic observations first to intermediate phonetic classes (phones/phonemes) and eventually to words. The normalization by the probability of the observation sequence, $P(O)$, can be omitted since it is constant for all W .

The acoustic observations take the form of multidimensional acoustic feature vectors that are extracted from the speech signal at equidistant time intervals, e.g., every 10 ms. Various feature extraction schemes have been devised; the most widely used representation is based on mel-frequency cepstral coefficients (MFCCs), which mimic the nonlinear frequency resolution characteristics of human auditory perception.

The classification of acoustic feature vectors is typically performed by hidden Markov models (HMMs). An HMM λ consists of a set of states $S = s_1, s_2, \dots, s_N$, a set of observation symbols $O = o_1, o_2, \dots, o_M$, and probability distributions governing the transitions between states and the emission of observation symbol from states. Given these parameters, the probability of an observation sequence $O = o_1, \dots, o_T$ given a model λ is computed as

$$P(O|\lambda) = \sum_S \prod_{t=1}^T a(s_t|s_{t-1})b(o_t|s_t) \quad (3)$$

where $a(s_t|s_{t-1})$ denotes the transition probability from state s_{t-1} to s_t , and $b(o_t|s_t)$ is the probability of the t th observation at state t . Thus, the global probability of an observation sequence given an HMM is defined as the product of all transition and observation probabilities over time points $1-T$, summed over all possible state sequences.

Of particular importance for the problem studied in this paper is the way in which the state-conditional observation probabilities, $b(o_t|s_t)$ in Eq. (3), are computed. In practice, the observation probabilities (as opposed to the transition probabilities) contribute most to the final probability score and are thus primarily responsible for acoustic-phonetic classification performance.

The most widespread model used in this context is a Gaussian mixture model, which computes the probability of the continuous acoustic observation vector \mathbf{o} at time t , $P(\mathbf{o}_t|s_j)$, as a weighted sum of M individual Gaussian probability density functions (mixture components):

$$p(\mathbf{o}_t|s_i(t)) = \sum_{m=1}^M c_{mi} \mathcal{N}(\mathbf{o}_t; \mu_{mi}, \Sigma_{mi}) \quad (4)$$

where μ_{mi} and Σ_{mi} are the mean vector and covariance matrix, respectively, of the m th mixture component of state i and c_{mi} is the mixture weight for that component. Each mixture component has the form of a Gaussian or Normal distribution:

$$\mathcal{N}(\mathbf{o}; \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} e^{-(1/2)(\mathbf{o}-\mu)' \Sigma^{-1} (\mathbf{o}-\mu)}. \quad (5)$$

where μ is the mean vector and Σ the covariance matrix of the distribution and d is the dimensionality of the feature space. The Gaussian mixture parameters, as well as the transition probabilities in the HMM, are estimated from training data using the expectation-maximization algorithm (Dempster *et al.*, 1977). Typically, HMMs are constructed for individual speech units such as phones, and are concatenated to form words, according to constraints specified in a pronunciation dictionary. The Viterbi algorithm is then used to find the best path through the concatenation of models.

B. Use of infant-directed training data in speech recognition

Given this background, two hypotheses can be made regarding the use of ID-style training data in speech recognition:

TABLE I. Cue words representing the vowels /i:/, /u:/, /ɒ/.

Vowel	Cue words		
/i:/	key	sheep	bead
/u:/	boot	shoe	spoon
/ɒ/	pot	top	sock

- (1) Phonetic classification using Gaussian mixture models is easier and more accurate when individual classes are well separated in the input space, i.e., when the class-conditional distributions overlap as little as possible. In previous experimental studies, the means of acoustic classes have been shown to be better separated in infant-directed than in adult-directed speech. While some researchers have also observed enlarged class variances in ID speech, all studies emphasize the strongly separated class means as the predominant property of ID speech. This suggests that the use of infant-directed training data might indeed be beneficial during training.
- (2) On the other hand, it is well known that matched training and test conditions are of utmost importance in ASR. When test data differ from training data, e.g., due to different recording conditions, presence of noise, or unknown accents, speech recognition performance deteriorates. For this reason, it might be hypothesized that training on infant-directed speech will not be useful for developing a speech recognizer for adult-directed speech, since the training and test conditions will be too different.

In order to test these hypotheses, separate speech recognizers are trained on two different data sets consisting either only of AD speech or only ID speech. Each of them is then tested on an AD data set and an ID data set, reflecting matched versus mismatched test conditions. In order to focus entirely on the contributions of the acoustic modeling component, we perform only isolated word recognition experiments—this serves to eliminate differences in performance that could be attributed to the language model or the vocabulary size (in ID speech, for instance, speakers tend to use a much smaller vocabulary).

III. DATA

The data used for the experiments described in this paper were drawn from a corpus of infant-directed speech provided by the Institute for Learning and Brain Sciences (formerly the Center for Mind, Brain and Learning) at the University of Washington. The corpus consists of 64 conversations by 32 different mothers. Each mother had two conversations, one with an adult experiment facilitator, the other with her infant. The data collection was designed to elicit certain cue words during both conversations, viz. nine monosyllabic words (shown in Table I) containing the vowels /i:/, /u:/, and /ɒ/ (the most distant points in the vowel triangle). Only these words were used for the present study since they had a sufficient number of samples and allow the comparison with previous studies focusing on the same vowels.

TABLE II. Distribution of training and test samples over words in AD and ID speech data sets.

Word	AD		ID	
	train	test	train	test
bead	136	26	196	56
boot	161	43	148	34
key	148	44	125	25
pot	121	25	136	29
sheep	138	41	148	45
shoe	216	54	154	35
sock	148	37	150	44
spoon	135	24	157	41
top	139	42	128	27

During the adult-infant conversations, toys of the same name (sheep, keys, etc.) were provided, which mothers used in interacting with their infants. In the adult-adult conversations, cue words were elicited simply by talking about related topics. Conversations were recorded using a far-field microphone (suspended from the ceiling near speakers' heads). For this reason, the recorded speech is far from studio-quality; rather, it reflects natural listening conditions. First, the use of a far-field rather than a close-talking microphone results in energy fluctuations in the speech signal due to the varying distance of the speaker's mouth to the microphone. Second, much of the speech is overlapped by background noise, such as infants crying or banging toys on the table, people entering and leaving the room, etc. Two infant-directed conversations were discarded *a priori* due to the presence of constant background noise. In addition, the adult-adult conversations contain many instances of speaker overlap or speech overlapped with laughter.

All audio files were time-segmented and transcribed orthographically by phonetically experienced transcribers. The time boundaries of noise events, speaker overlaps, and all instances of cue words were marked. These time marks were double-checked in a second transcription pass, and all instances of cue words were extracted from the audio data. Each instance was categorized by a trained listener with respect to the degree of background noise (on a scale from 1 to 10, with 1 being the lowest and 10 being the highest noise level), and with respect to the presence or absence of pitch-accent on the word (on a scale from 1 to 5, with 1 being the lowest confidence and 5 being the highest). The latter was determined by listening to the word in its context in the original speech file. All samples of cue words that had a noise level higher than three were discarded. Furthermore, only data from those speakers with sufficient instances in both the ID and the AD conversations were selected. As a result, the data from eight speakers could not be used at all, such that the final data set used for the experiments reported below consists of the cue words from 22 different speakers. The remaining samples were randomly assigned to training and test sets. For each speaking mode (ID and AD), there were 1342 training and 336 test samples; every speaker had the same number of training and test samples in both the AD and ID data sets. Table II shows the distributions of samples over training and test sets. Further balancing of the data in

TABLE III. Word accuracy (%) for speech recognizers trained on AD versus ID speech and tested on matched versus mismatched conditions.

		Test	
		AD	ID
train	AD	95.5	81.6
	ID	90.2	93.5

terms of speaker and word identities (e.g., enforcing exactly the same number of samples for each word) would have resulted in sets with too few training samples. Although the data sets are very small by ASR standards, they are larger than the samples used by most other phonetic studies on infant-directed speech.

IV. EXPERIMENTS AND RESULTS

The original recorded data files were sampled at 16 kHz and preprocessed using MFCC analysis. The acoustic front-end consisted of 39 coefficients (12 MFCCs, normalized log energy, and their first and second derivatives), extracted every 10 ms with a window size of 25 ms. Cepstral mean and variance normalization were applied on a per-conversation basis to reduce the effects of recording conditions and speaker variation. Whole-word acoustic HMMs were then constructed for each cue word, with eight emitting states and two Gaussian mixture components per state. In accordance with standard practice in small-vocabulary ASR, whole-word- rather than subword-unit-based HMMs were used in order to achieve better coarticulation modeling. The models had a left-to-right topology without skip transitions, e.g., each state could transition only to itself or the immediately following state; states could not be skipped. Given that most ASR systems use three-state models for phone units (which are on average 30 ms long), the number of states reflects the expected duration of words consisting of three phones. Diagonal covariance matrices were used in each mixture component. The models were trained using three iterations of EM; recognition was performed by one-best Viterbi search. Due to the small size of the data set, more advanced modeling techniques which increase the number of parameters (such as context-dependent models or full covariance matrices) were not used.

A. Experiment I: Baseline study

The performance of the recognizers is measured as word accuracy, i.e., the percentage of correctly recognized samples in the test set. The results (Table III) show that absolute recognition performance is highest on the AD data set when using an AD-trained recognizer (upper left corner of Table III). In both conditions, the recognizer trained and tested on matched sets (the diagonal of the table) performs better than the corresponding recognizer trained on mismatched data. However, systems behave differently with respect to the relative degradation of recognition performance under mismatched conditions. The relative degradation of the AD-trained system on ID speech is 16.5% whereas the ID-trained system applied to AD speech shows a relative loss of 3.3%

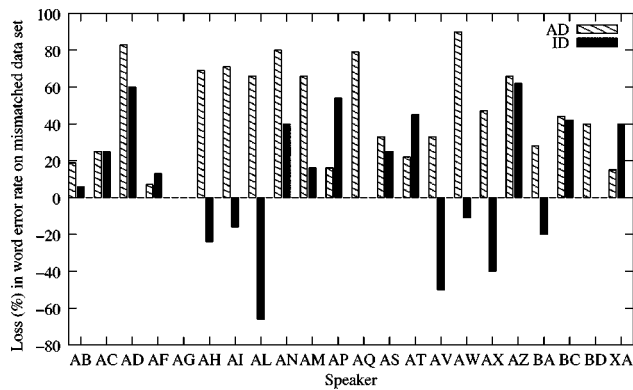


FIG. 1. Per-speaker relative degradation (in % word error rate) of AD and ID recognizers on mismatched test data.

accuracy. This difference was statistically significant at the 0.0001 level, using a difference of proportions significance test. The absolute differences in accuracy are on the same scale (loss of 13.9% absolute of AD recognizer on ID speech and 3.3% of the ID recognizer on AD speech).

The relative degradation was also computed on a per-speaker basis. Results are displayed in Fig. 1. For 16 out of 22 speakers, the AD-trained recognizer showed a stronger degradation on ID speech (indicated by the striped bars) than vice versa (black bars). For two speakers, the performance did not change, and four speakers showed a stronger relative degradation on AD than on ID speech. The ID-trained recognizers also showed a better performance on the AD data set than on the ID test set for several speakers (the bars extending into the negative region in Fig. 1). From this we can conclude that the ID-trained recognizers perform more robustly under mismatched training and test conditions. There was no correlation with the amount of training or test data used per speaker.

B. Analysis

An analysis of the recognition errors showed that confusions occurred mainly between words sharing the same vowel (e.g., confusions between *key* and *bead* or *top* and *sock*) and words of the /u:/ and /ɒ/ categories (such as *spoon* and *sock*). Confusions between /i:/ words and the other categories were rare, although some examples of *shoe*–*sheep* confusions did occur, possibly due to strong coarticulation of the /u:/ vowel with the palatal fricative /sh/.

For a more detailed analysis it was necessary to look at individual phonetic segments. In order to automatically extract signal portions corresponding to subword units, we trained individual phone models, using a transcription of words in terms of their constituent phones. The preprocessing of the signals was identical to the recognition experiments described above. The phone models consisted of three emitting states with two Gaussian mixture components each; the topology was left-to-right without any skip transcriptions. The initial model parameters were set to the global mean and variance of the data, i.e., the parameters were computed from all feature vectors, regardless of phonetic class. The models were trained in three iterations of EM and were then used in a forced alignment procedure in order to obtain

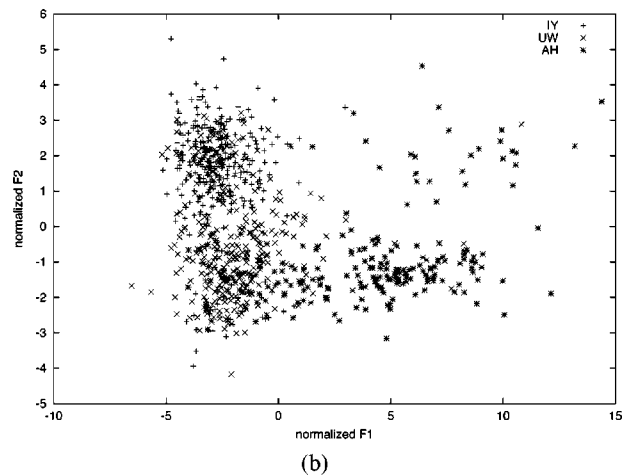
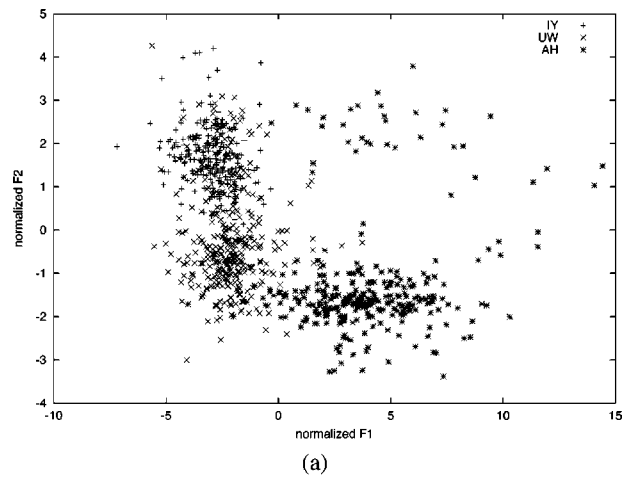


FIG. 2. Distribution of vowel classes in normalized formant space in AD speech (a) and ID speech (b).

time boundaries of the constituent phones. During this procedure, the reference transcription was provided to the recognizer, and the Viterbi algorithm was used to identify the globally best time alignment of the model sequence defined by the transcription. The resulting time alignments were then checked for accuracy and the corresponding time segments were extracted. This data was used to analyze the distribution and separability of phonetic classes under different acoustic representations and prosodic conditions, as described in the following sections.

1. Separability of vowel classes

Standard formant-based analysis was applied to the vowel spaces in AD and ID speech. For each vowel segment (as identified by the forced alignment), formant values were obtained using the formant program of the Entropics ESPS package, and they were averaged over the entire length of the vowel. Cross-speaker normalization in the form of mean and variance normalization was applied.

The class distributions in the space defined by the normalized first two formants are shown in Fig. 2. Similar to previous studies which have investigated phonetic discriminability in either speech perception or automatic classification scenarios (e.g., Nossair and Zahorian, 1991; Zahorian and Jagharghi, 1993; Hillenbrand *et al.*, 1995), we computed

TABLE IV. Comparison of within-class distance (V^2), within-class distance (D^2), Fisher ratio (\mathcal{F}) (the ratio of within-class to between-class distance), and average Euclidean distance (E) between class means for AD and ID vowel data. The upper row shows results based on speaker-normalized formant values; the lower row shows results for PCA-mapped mel-frequency cepstral coefficients.

	AD				ID			
	V^2	D^2	\mathcal{F}	E	V^2	D^2	\mathcal{F}	E
Formants	4.98	24.32	0.17	3.96	8.73	20.94	0.29	3.84
MFCCs	1.08	12.72	0.08	3.24	2.70	8.82	0.23	2.89

various numerical measures of class separation. One of these is the Fisher ratio \mathcal{F} (e.g., Schuermann 1996) defined as

$$\mathcal{F} = \frac{V^2}{V^2 + D^2}, \quad (6)$$

where

$$V^2 = \sum_{k=1}^K P_k \text{trace}[\Sigma_k] \quad (7)$$

and

$$D^2 = \frac{1}{1 - \sum_{k=1}^K P_k^2} \sum_{k=1}^K \sum_{j=1}^K P_k P_j (\mu_k - \mu_j)^2 \quad (8)$$

where K is the number of classes, Σ_k is the covariance matrix of the k th class, μ_k is the mean of the k th class, and P_k is the prior probability of the k th class. V^2 measures the within-class variance of the features with respect to the class mean. D^2 denotes the interclass distance, i.e., the distance between class means, weighted by the class priors. \mathcal{F} thus expresses the ratio of within-class distance to the between-class distance and ranges between 0 and 1. A lower \mathcal{F} value indicates better separability. We also computed the average Euclidean distance between the three class means. These measures are shown in Table IV.

The results indicate a slightly better separation of class means in AD speech than in ID speech. Overall class separability is noticeably poorer in ID speech than in AD speech; this is due to a stronger degree of class overlap caused primarily by larger class variances.

In order to eliminate possible effects caused by the automatic segmentation, the analysis was repeated for average formant values computed only over the center third of each vowel. However, no significant difference was found between the results.

Formant values are useful for comparing the present analysis to other phonetic studies of vowel spaces; however, they are not used in the speech recognizers. It is not immediately obvious whether class distributions in formant space can be equated with class distributions in the 39-dimensional MFCC space that defines the recognizer front-end. The above analysis was therefore repeated for the MFCC data. Principal components analysis (PCA) (see, e.g., Duda *et al.*, 2001) was applied to the data and feature vectors, and the feature space was mapped to the coordinate system defined by the first two principal components, in order to be able to display the result graphically. Table IV lists the Fisher ratio

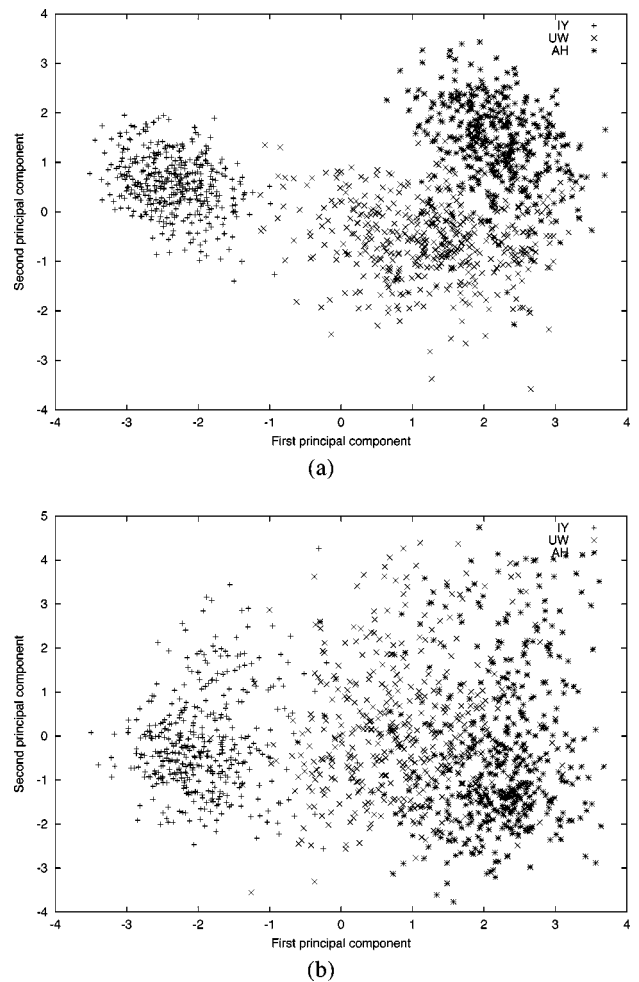


FIG. 3. Distribution of vowel classes in MFCC space mapped to the first two principal components, in AD speech (a) and ID speech (b).

and distance measures computed on this data, and Fig. 3 shows the corresponding distributions. Similar results can be observed: the average distance of class means is slightly larger and class separation is more pronounced in AD speech. Class separability measures computed on the full 39-dimensional feature space showed the same pattern.

One factor contributing to the variability in ID speech may be that no distinction was made between vowels that are typical examples of “motherese,” i.e., vowels that exhibit strong hyperarticulation, and those that are not. Our definition of “infant-directed speech” includes the ensemble of speech used in infant-directed conversations. Although this definition reflects natural listening conditions more faithfully, ID speech is a continuum, and a further distinction can be made between different points within this continuum. To this end, pitch-accented versus non-pitch-accented vowels were considered separately. Vowels bearing linguistic stress or pitch-accent are often articulated more clearly, whereas unstressed or unaccented vowels show a stronger degree of coarticulatory reduction, with a concomitant decrease in the size of the vowel space. This has been verified in a number of experimental and corpus-based studies (e.g., Lindblom, 1963; Sluijter and van Heuven, 1996; Kirchhoff and Bilmes, 1999). Vowels that are more representative of ID speech can be expected to be perceived as pitch-accented. We extracted

TABLE V. Within-class distance (V^2), between-class distance (D^2), Fisher ratio (\mathcal{F}) (the ratio of within-class to between-class distance), and average Euclidean distance (E) between class means for pitch-accented vowels in AD and ID speech.

AD				ID			
V^2	D^2	\mathcal{F}	E	V^2	D^2	\mathcal{F}	E
1.91	8.39	0.19	2.86	3.01	8.11	0.27	2.91

pitch-accented samples according to the manual annotations described above (Sec. II). For each condition, we extracted the 220 vowel samples with the highest pitch-accent values. The analysis measures were then recomputed for these data sets based on PCA-mapped MFCCs. The results are shown in Table V.

In spite of the elimination of non-pitch-accented vowel samples, the pattern observed before still holds: class overlap is greater in ID speech than in AD speech. However, the average mean Euclidean distance is now slightly greater than in ID than in AD speech. These observations confirm results obtained in previous studies (see above Sec. I). Although studies focusing on content words mostly report an expansion of the vowel space, van de Weijer (2001) explicitly compared vowels in content versus function words and found that expansion of the vowel space in ID speech occurred in content words but was reversed in function words, which are presumably not pitch-accented.

2. Separability of consonant classes

Most previous studies of infant-directed speech have focused on the acoustic properties of vowels. Consonants, on the other hand, are less well studied. An exception is Sundberg and Lacerda (1999) and Sundberg (2001), where the effects of infant-directed speech on the voice onset time (VOT) of stop consonants was studied. It was shown that VOT is significantly shorter in infant-directed speech than AD speech for infants in younger infants (three months) but changed to longer, overspecified patterns in older infants (11–14 months).

The error patterns in the recognizer output indicate that word confusions were caused not only by poor separation of vowels, but also by consonant substitutions. Therefore, the class separability measures described above were also applied to consonant distinctions, viz. the stops /b/ vs. /p/ in syllable-initial position, and the fricatives /sh/ vs. /s/ in syllable-initial position. These pairs were chosen because they represent the same broad phonetic class in comparable prosodic contexts; in particular, syllable-initial consonants were chosen because they tend to be less strongly coarticulated than syllable-final consonants. As before, the MFCC space was mapped to two dimensions using PCA. Table VI shows the resulting measurements. Whereas the average class mean distance is fairly similar, the Fisher ratio is again lower in AD speech.

C. Experiment II: Linear discriminant analysis

In pattern recognition, a common technique to enhance acoustic class separability is linear discriminant analysis

TABLE VI. Between-class distance (D^2), within-class distance (V^2), Fisher ratio (\mathcal{F}), and average Euclidean distance (E) between class means for syllable-initial consonant pairs /b/-/p/ and /sh/-/s/.

	AD				ID			
	V^2	D^2	\mathcal{F}	E	V^2	D^2	\mathcal{F}	E
/b/-/p/	1.83	6.83	0.21	1.06	2.05	4.99	0.29	1.04
/sh/-/s/	0.77	3.24	0.19	0.83	1.21	1.63	0.43	0.82

(LDA) (see, e.g., Duda *et al.*, 2001). LDA is a transformation designed to maximize the between-class distance while minimizing within-class distance; it is thus related to the Fisher criterion discussed above. Applying LDA to the infant-directed speech data should reduce the class overlap and produce better recognition results, possibly with the effect of smoothing out differences in performance between the AD and ID recognizer. We used LDA on both the AD and ID data sets. In each case, the transformation parameters were estimated on the training set and were then applied to both the training and the test set. The transformed data was used to train the speech recognizers, using exactly the same modeling procedures as before (see above Sec. IV A). The recognition results, listed in Table VII., show that recognition performance improves for both recognizers under both conditions. The relative improvement compared to the baseline results (Table III) is greatest for the AD-trained recognizer on ID test conditions. As before, however, the relative loss in word accuracy on the mismatched condition is more severe for the AD-trained recognizer than for the ID-trained recognizer; the difference is statistically significant at the 0.001 level.

V. DISCUSSION

Automatic speech recognizers trained on ID and AD speech, respectively, were applied to both matched and mismatched test sets. It was found that, on average, matched conditions produced better results than mismatched conditions, confirming the second of the hypotheses stated in Sec. II B. ID-trained recognizers performed better on AD speech test sets for some but not for most most speakers. However, the relative degradation of ID-trained recognizers on AD speech was significantly less severe than in the reverse case.

An analysis of a subset of phonetic class distributions in ID and AD speech showed that ID speech was characterized by a stronger class overlap, which provides an explanation of the recognition results: models trained on strongly overlapped classes can accommodate well-separated test data, but models trained on well-separated data will fail to correctly

TABLE VII. Word accuracy (%) of AD- and ID-trained speech recognizers under identical and mismatched test conditions, after application of LDA transformations.

		test	
		AD	ID
train	AD	97.4	91.8
	ID	92.9	94.7

classify test samples whose distributions show a stronger overlap. Samples in the intersection of the class decision boundaries may receive equally likely scores from the acoustic models, thus causing more confusions. The use of linear discriminant analysis to enhance class separation improved recognition results in all test conditions.

The automatic classification procedure used here does not claim to be a cognitively adequate model of infant perceptual learning. First, automatic speech recognizers are trained on a limited, well-defined training set, whereas it is, in general, impossible to quantify how much training data infants have previously been exposed to. Data collections where the amount of training data for a particular phonetic contrast can be controlled precisely [as in Kuhl *et al.*, (2003)] might turn out to be an interesting testbed for pattern recognition algorithms. Second, the acoustic representation and/or the statistical modeling techniques used in ASR systems certainly have deficiencies compared to human speech perception, since ASR performance still falls short of human performance. Nevertheless, the analyses presented above highlight interesting research questions for the study of infant speech perception. Previous studies on ID speech have emphasized the greater distance between vowel class means and have concluded that “motherese” might facilitate phonetic category learning. In contrast, it was found here (using comparatively large sets of samples and speakers) that class separability is actually poorer in ID speech. A possible reason is that *all* samples produced from infant-directed conversations were included, whereas previous studies may have focused on samples that are most “motherese-like” and thus occupy extreme positions in the vowel space. The slightly greater separation of the class means of pitch-accented vowels observed above confirms this interpretation. Nevertheless, the stronger class overlap is still present in those sets, suggesting that ID speech is poor training data. This result is in line with earlier studies comparing the intelligibility of speech directed to adults versus speech directed to children of 1–3 years (Bard and Anderson, 1983, 1994). In those studies, isolated words excised from either adult-adult or adult-child conversations had to be identified by adult listeners and young children. In Bard and Anderson (1994) additional context was provided for the tokens. It was found that word intelligibility was inversely related to word predictability: predictable words without their context were actually less intelligible in child-directed than in adult-directed speech. The conclusion drawn from this study was that adults might reduce predictable words more when talking to their infants/children than when talking to adults in order to draw the child’s attention to referents that are new in the discourse or in the extralinguistic environment. The observation made by van de Weijer (2001) that ID speech exhibits an enlarged vowel space in content words but a reduced vowel space in function words (see Sec. I) supports this analysis. It seems plausible that words which are predictable or already given in the discourse are underarticulated in favor of words that are new, some of which may then exhibit the typical motherese effect.

Our observations also pose the question of why infants are able to acquire phonetic categories and generalize to new

test samples in spite of strongly overlapped training data—we may assume that the majority of everyday speech that infants are exposed to consists of predictable function words [see also Cutler (1993)]. Several explanations may be advanced, e.g., that auditory representations in human speech perception are more invariant, that different (distribution-free) classification mechanisms are involved, or that listeners perform some form of data normalization, selection, or variance reduction, similar to the LDA transformation. Not all of these explanations are equally likely—it has been shown experimentally that infants do respond to distributions in auditory training data (Maye *et al.*, 2002). Moreover, although the acoustic representations and normalization techniques used in this study may not model human auditory representations perfectly, it is safe to assume that they are correlated. It seems most plausible that infants make use of perceptual variance reduction techniques. One possible candidate for such a technique is the perceptual magnet effect (Kuhl, 1991; Kuhl *et al.*, 1992). This effect is often described as a perceptual warping in the sense that auditory stimuli close to a “prototypical” representation of the category cannot be distinguished perceptually, thus decreasing within-class variance. Although the existence of this effect as well as the precise nature of the magnet theory have been under much discussion (Lively and Pisoni, 1997; Lotto *et al.*, 1998; Guenther, 2000; Lotto, 2000), it is striking that phonetic category perception does not occur until fairly late in the first year of life, after the emergence of the magnet effect as dated by proponents of this theory. It is unlikely that only the *amount* of speech data encountered by infants is responsible for the onset of categorical perception, or that categories are formed simply by passively processing all available input without some active contribution by the perceptual system: naturally occurring speech data is extremely variable in its distribution, and, as the present study shows, may even be more variable in ID speech than in AD speech. It may be assumed that infants selectively filter the available input, or perform a compaction of stored class representations [see recent evidence from neural imaging (Guenther and Bohland, 2002)]. Filtering might be done by focusing on the most motherese-like items in the speech input, which then become the prototypes of phonetic classes.

In summary, two conclusions can be drawn from this study. First, from an ASR point of view, ID-style speech seems to be of no immediate benefit to present-day ASR systems: matched training data still leads to superior results on AD test data. Furthermore, filtering of the training data in the way suggested above would not be advisable for ASR systems because of the resulting reduction in training material. The second conclusion is that parameters encoding contextual information (either linguistic or extralinguistic) should be explicitly integrated into data selection methods in acoustic-phonetic studies of ID speech, and into actual theories of infant speech perception. It should be noted that the ASR framework provides a tool for quantifying linguistic predictability (in the form of language model perplexity), which can in turn be used in data selection. It thus seems that the main benefits of combining ASR technology and the study of speech perception lie in the use of ASR techniques

for speech data selection and statistical representation as well as for building and testing perceptual models.

ACKNOWLEDGMENTS

This work was funded by the Institute for Learning and Brain Sciences, University of Washington. We would like to thank Patricia Kuhl and Bart de Boer for providing the speech data and three anonymous reviewers for their comments.

- Andruski, J. E., and Kuhl, P. K. (1996). "The acoustic structure of vowels in mothers' speech to infants and adults," in *Proceedings of ICSLP*, pp. 1541–1544.
- Bard, E. G., and Anderson, A. H. (1983). "The unintelligibility of speech to children," *J. Child Lang* **10**, 265–292.
- Bard, E. G., and Anderson, A. H. (1994). "The unintelligibility of speech to children: effects of referent availability," *J. Child Lang* **21**, 623–648.
- Burnham, D., Kitamura, C., and Vollmer-Conna, U. (2002). "What's new, pussycat? on talking to babies and animals," *Science* **296**, 1435.
- Cooper, R. P., and Aslin, R. N. (1994). "Development differences in infant attention to the spectral properties of infant-directed speech," *Child Dev.* **65**, 1663–1677.
- Cutler, A. (1993). "Phonological cues to open- and closed-class words in the processing of spoken sentences," *J. Psycholinguistic Res.* **22**, 109–131.
- Damper, R. I., and Hanard, S. R. (2000). "Neural network models of categorical perception," *Percept. Psychophys.* **62**(4), 843–867.
- de Boer, B., and Kuhl, P. K. (2003). "Investigating the role of infant-directed speech with a computer model," *ARLO* **4**, 129–134.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). "Maximum-likelihood from incomplete data via EM algorithm," *J. R. Stat. Soc. Ser. B. Methodol.* **39**(1), 1–38.
- Duda, R. O., Hart, P. E., and Stork, D. G. (2001). *Pattern Classification*, 2nd edition (Wiley, New York).
- Fernald, A., and Kuhl, P. K. (1987). "Acoustic determinants of infant preference for motherese speech," *Inf. Behav. Dev.* **10**, 279–293.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., Boysson-Bardies, B., and Fukui, I. (1989). "A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants," *J. Child Lang* **16**, 477–501.
- Grieser, D. L., and Kuhl, P. K. (1988). "Maternal speech to infants in a tonal language: support for universal prosodic features in motherese," *Dev. Psychol.* **24**, 14–20.
- Guenther, F. H. (2000). "An analytical error invalidates the 'depolarization' of the perceptual magnet effect," *J. Acoust. Soc. Am.* **107**, 3576–3580.
- Guenther, F. H., and Bohland, J. W. (2002). "Learning sound categories: a neural model and supporting experiments," *Acoust. Sci. Technol.* **23**(4), 213–220.
- Hazan, V., and Simpson, A. (2000). "The effect of cue-enhancement on consonant intelligibility in noise: speaker and listener effects," *Lang Speech* **43**, 273–294.
- Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (1995). "Acoustic characteristics of American English vowels," *J. Acoust. Soc. Am.* **97**, 3099–3111.
- Holt, L. L., Lotto, A. J., and Kluender, K. R. (1998). "Incorporating principles of general learning in theories of language acquisition," M. Gruber, C. D. Higgins, K. S. Olson, and T. Wysocki (eds.) in *Chicago Linguistic Society: Papers from the Panels*, 34(2), pp. 253–268, Chicago: Chicago Linguistic Society.
- Jusczyk, P. W., Luce, P. A., and Charles-Luce, J. (1994). "Infants' sensitivity to phonotactic patterns in the native language," *J. Mem. Lang.* **33**, 630–645.
- Kirchhoff, K., and Bilmes, J. (1999). "Statistical acoustic effects of coarticulation," in *Proceedings of the 14th Int. Congress on Phonetic Sciences*, San Francisco, CA.
- Kuhl, P. K. (1991). "Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not," *Percept. Psychophys.* **50**, 93–107.
- Kuhl, P. K., Tsao, F. M., and Liu, H. M. (2003). "Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning," *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9096–9101.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (1992). "Linguistic experience alters phonetic perception in infants by 6 months of age," *Science* **255**, 606–608.
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., and Lacerda, F. (1997). "Cross-language analysis of phonetic units in language addressed to infants," *Science* **277**, 684–686.
- Lindblom, B. (1963). "Spectrographic study of vowel reduction," *J. Acoust. Soc. Am.* **35**, 1773–1781.
- Liu, H. M., Kuhl, P. K., and Tsao, F. M. (2003). "An association between mothers' speech clarity and infants' speech discrimination skills," *Deve. Sci.* **6**(3), F1–F10.
- Lively, S. E., and Pisoni, D. B. (1997). "On prototypes and phonetic categories: a critical assessment of the perceptual magnet effect in speech perception," *J. Exp. Psychol. Hum. Percept. Perform.* **23**(6), 1665–1679.
- Lotto, A. J. (2000). "Reply to An analytical error invalidates the 'depolarization' of the perceptual magnet effect," *J. Acoust. Soc. Am.* **107**, 3578–3581.
- Lotto, A., Kluender, K., and Holt, L. L. (1998). "Depolarizing the perceptual magnet effect," *J. Acoust. Soc. Am.* **103**, 3648–3655.
- Maye, J., Werker, J. F., and Gerken, L. A. (2002). "Infant sensitivity to distributional information can affect phonetic discrimination," *Cognition* **82**, B101–B111.
- Moore, R. K. (2003). "A comparison of the data requirements of automatic speech recognition systems and human listeners," in *Proceedings of Eurospeech*, pp. 2581–2584.
- Nearey, T. M. (1997). "Speech perception as pattern recognition," *J. Acoust. Soc. Am.* **101**, 3241–3254.
- Nossair, Z. B., and Zahorian, S. A. (1991). "Dynamic spectral shape features as acoustic correlates for initial stop consonants," *J. Acoust. Soc. Am.* **89**, 2978–2990.
- Payton, K., Uchanski, R., and Braidia, L. (1994). "Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing," *J. Acoust. Soc. Am.* **95**, 1581–1592.
- Picheny, M., Durlach, N., and Braidia, L. (1986). "Speaking clearly for the hard of hearing II: acoustic characteristics of clean and conversational speech," *J. Speech Hear. Res.* **32**, 93–103.
- Protopapas, A. (1999). "Connectionist modeling of speech perception," *Psychol. Bull.* **125**(4), 410–436.
- Protopapas, A., and Calhoun, B. (2000). "Adaptive phonetic training for second language learners," in *Proceedings of InSTIL (2nd International Workshop on Integrating Speech Technology in Language Learning)*, Dundee, UK, pp. 31–38.
- Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). "Statistical learning by 8-month-olds," *Science* **274**, 1926–1928.
- Scharenborg, O., Boves, L., and de Veth, J. (2002). "ASR in a human word recognition model: generating phonemic input for Shortlist," in *Proceedings of ICSLP*, pp. 633–636.
- Scharenborg, O., McQueen, J. M., ten Bosch, L., and Norris, D. (2003). "Modelling human speech recognition using automatic speech recognition paradigms in SpeM," in *Proceedings of Eurospeech*, pp. 2097–2100.
- Schuermann, J. (1996). *Pattern Classification: A Unified View of Statistical and Neural Approaches* (Wiley, New York).
- Sluijter, A. M. C., and van Heuven, V. J. (1996). "Acoustic correlates of linguistic stress and accent in Dutch and American English," in *Proceedings of ICSLP*, pp. 630–633.
- Sundberg, U. (2001). "Consonant specification in infant-directed speech. Some preliminary results from a study of voice onset time in speech to one-year-olds," Lund University Department of Linguistics Working Papers **49**, 148–151.
- Sundberg, U., and Lacerda, F. (1999). "Voice onset time in speech to infants and adults," *Phonetica* **56**, 186–199.
- van de Weijer, J. (2001). "Vowels in infant- and adult-directed speech," Lund University Department of Linguistics Working Papers **49**, 172–175.
- Zahorian, S. A., and Jagharghi, A. J. (1993). "Spectral-shape features versus formants as acoustic correlates for vowels," *J. Acoust. Soc. Am.* **94**, 1966–1982.

The acoustics of Japanese wooden drums called “mokugyo”^{a)}

Masahiro Sunohara and Kenji Furihata^{b)}

Department of Electrical and Electronic Engineering, Faculty of Engineering, Shinshu University, 4-17-1 Wakasato, Nagano, 380-8533 Japan

David K. Asano

Department of Information Engineering, Faculty of Engineering, Shinshu University, 4-17-1 Wakasato, Nagano, 380-8533 Japan

Takesaburo Yanagisawa

Faculty of Engineering, Shinshu University, 4-17-1 Wakasato, Nagano, 380-8533 Japan

Atsuyoshi Yuasa

Mokugyokan, 6,282 Setogawa, Ogawamura, 381-3304 Japan

(Received 20 July 2004; revised 14 January 2005; accepted 19 January 2005)

A drumlike traditional Japanese instrument, the mokugyo, is experimentally discussed. First, the acoustic characteristics of 176 mokugyos with diameters ranging from 7.5 to 120 cm and three drumsticks were measured. Results show that (a) the sound spectra consist of two common peaks [F_1 (Hz): first peak frequency, F_2 (Hz): second peak frequency] close together, with an average ratio (F_2/F_1) of 1.15, and (b) a drumstick beating the mokugyo is translated into an impact force applied over a period of time from 1 to 6 ms related to the mass and stiffness of the material wrapped around the tip of the drumstick. Second, to evaluate the acoustic response of a mokugyo in the final tuning process, the mechanical and acoustical analogy between the mokugyo and a bass reflex loudspeaker is theoretically and experimentally discussed. Results show that the model can be estimated within a relative error of 0.52% from the mass of wood chips. Finally, from a psychological experiment, the timbre of the mokugyo shows higher scores on psychological scales when the ratio (F_2/F_1) becomes 1.15. © 2005 Acoustical Society of America.

[DOI: 10.1121/1.1868192]

PACS numbers: 43.75.Kk [NHF]

Pages: 2247–2258

I. INTRODUCTION

In most of the rituals for Buddhism, the sutras are recited to the rhythm created by several percussion instruments such as a mokugyo, a keisu, and a large bowl-shaped gong.¹ A mokugyo is a round drum made of wood used during Zen rituals like the recitation of the sutras. Buddhists recite sutras while striking a mokugyo with a regular rhythm. Most Japanese have been familiar with the mokugyo sound since childhood. The word, “mokugyo,” means “a wooden fish” in Japanese. In Buddhism the fish, which never sleeps, symbolizes wakefulness. It was introduced from China to Japan around the 17th century. From an acoustic engineering standpoint, the mokugyo has never been investigated.

In recent years, physical modeling of the sound production mechanism of percussion instruments like a mokugyo has been studied. For example, Chaigne *et al.*^{2,3} proposed a physical model of xylophone sound production and synthesized xylophone sounds based on the proposed model. They reported that the synthesized sound was close to the real sound of the xylophone, and that the quality of the synthe-

sized sound could be controlled by adjusting the physical structure of the xylophone using the proposed model. L. Rhaoui *et al.*⁴ also proposed a physical model of the kettle-drum and its synthesized sounds from a similar standpoint. On the other hand, Obata *et al.*⁵ investigated acoustic characteristics of a Japanese traditional percussion instrument, the “ôdaiko.” They reported that the fundamental frequency component of ôdaiko sounds involved approximately 25 beats per second. The beats were produced by a coupled vibration between the drumskins on both sides of the body and the enclosed air. In the same way, Ando *et al.*^{6,7} investigated and synthesized the sound of the “tudumi,” and Aoki *et al.*⁸ investigated that of the “washô.”

Figure 1(a) shows the front view of a general mokugyo with a 37.5-cm diameter, (b) shows its side view with a 32.0-cm height, and (c) shows a drumstick with a 43.0-cm length. The body is made of camphorwood with a density of 0.16 g/cm³. Reliefs of dragons or fishes are carved on the body. The sound is produced by striking the contact area with a drumstick. As indicated by the dotted line in Fig. 1(b), the body is carved from a single block of wood to have a large internal cavity, which is quite a feat, given its narrow slit opening (intended to resemble the mouth of a fish). The cavity walls are 2.5 cm thick on average. These components form a Helmholtz resonator system. The handle of the drumstick is made of wood, and the tip is wrapped in rubber.

^{a)}Portions of this work were presented in “Acoustic characteristics of Japanese wooden-drum such as mokugyo,” 17th International Congress on Acoustics, Roma, Italy, 5A.16.02, September 2001, and “Timbre of the Mokugyo of Japan,” 17th International Congress on Acoustics, Roma, Italy, 5A.16.03, September 2001.

^{b)}Author to whom correspondence should be addressed. Electronic-mail: kennfur@gipwc.shinshu-u.ac.jp

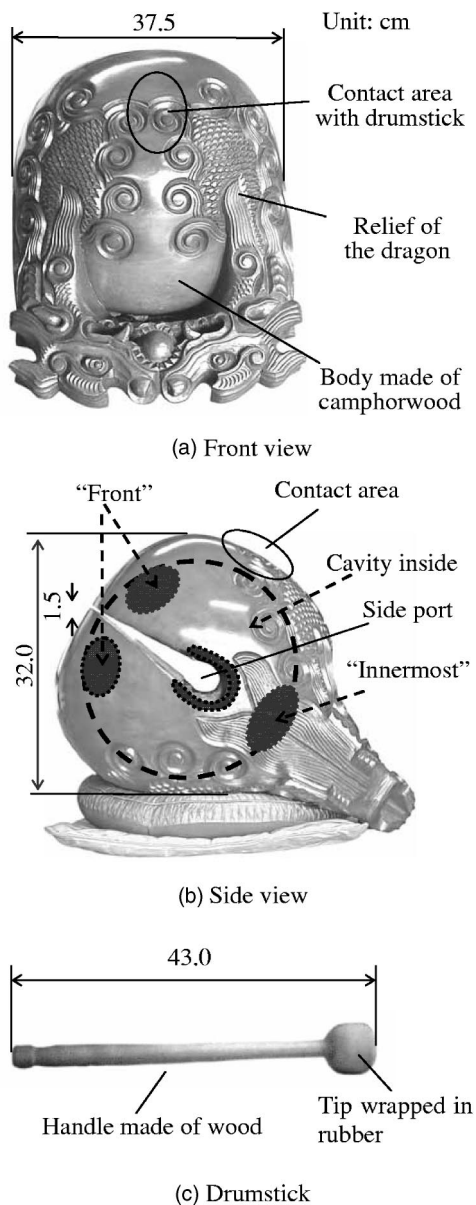


FIG. 1. The mokugyo. (a) Front view. (b) Side view. (c) The drumstick.

Typical drumstick tips are often wrapped in thread, cloth, or leather. The kinds of wrapping materials may be an important factor for producing beautiful timbre, since the timbre varies with the material. Therefore, the artisan making a mokugyo also makes a drumstick exclusively for the mokugyo.

Traditional skills for making mokugyo are endangered, since the number of artisans has decreased. One of the causes is the technical difficulty involved. The most difficult process is the adjustment of the timbre. This is done by chiseling the cavity of the side port little by little based only on the artisan's intuition. This process suggests that the physical characteristics which determine the timbre are affected by various structural elements of the body. If it is possible to control such physical parameters numerically using a physical model, the desired timbre could be produced easily.

In this paper, acoustic characteristics of mokugyos are investigated experimentally, theoretically, and psychoacoustically. In Sec. II, five physical characteristics of the

mokugyo and its drumstick are measured and discussed: mokugyo sound spectrum, vibration of the body, waveform and spectrum of the driving forces of the drumstick, directivity, and sound intensity vector of the acoustic radiation of the mokugyo. In Sec. III, a physical model of the mokugyo is proposed in which it is assumed that the mechanical vibratory section of the body and the Helmholtz resonator system are coupled. The mokugyo sound is simulated in the frequency domain. Parameter values used in the proposed model are estimated by comparing the spectrum of the simulated sound with that of the original sound. In Sec. IV, the proposed model is evaluated. Three areas in the cavity, which are called "front," "innermost," and "side port," are chiseled, and then the difference in the spectrum before and after the chiseling is measured and evaluated. In Sec. V, the relation between favorable timbre and physical characteristics of the mokugyo is experimentally examined with the semantic differential method in order to establish indices for evaluating mokugyo timbre.

II. EXPERIMENTS

In order to investigate the physical and acoustic characteristics of the mokugyo, five characteristics were measured: sound spectrum, vibration of the body, drumstick driving force, directivity, and sound intensity vector.

A. Sound spectrum

1. Experimental method

One hundred seventy-six mokugyos were used for this measurement. The minimum diameter of the mokugyos was 7.5 cm and the maximum was 120 cm with an average diameter of 46 cm. Each mokugyo was struck with a drumstick at the contact area shown in Fig. 1. This experiment was conducted in an anechoic room with a volume of 60 m³. To record the sounds a microphone was set at a distance of 100 cm away from the side port, and the output signals were recorded on DAT (Technics: SV-MD11) with a sampling frequency of 48 000 Hz. The recorded sounds were analyzed using a (fast Fourier transform) (FFT) with a rectangular window of 32 768 samples.

2. Results

Figure 2 shows waveforms and spectra for three mokugyos with typical characteristics. Time scales in the figures are different from each other. The rising edge of the waveform (i.e., when the drumstick contacts the body) is set to "0" in the time scale. The beats were observed for all three waveforms whose periods are (a) 7.2 ms for the mokugyo with a diameter of 9 cm, (b) 46 ms for that with a diameter of 37 cm, and (c) 110 ms for that with a diameter of 90 cm. Three peaks were observed in all sound spectra. The two peaks at lower frequencies are significant. In this paper, the three peaks are defined as the first, second, and third resonance frequencies, F_1 , F_2 , and F_3 (Hz) in increasing order. Each resonance frequency decreases as the diameter increases, and F_1 and F_2 are close. These two close peaks are a general physical characteristic for the sound of the mokugyo, since such peaks were observed for 174 out of 176

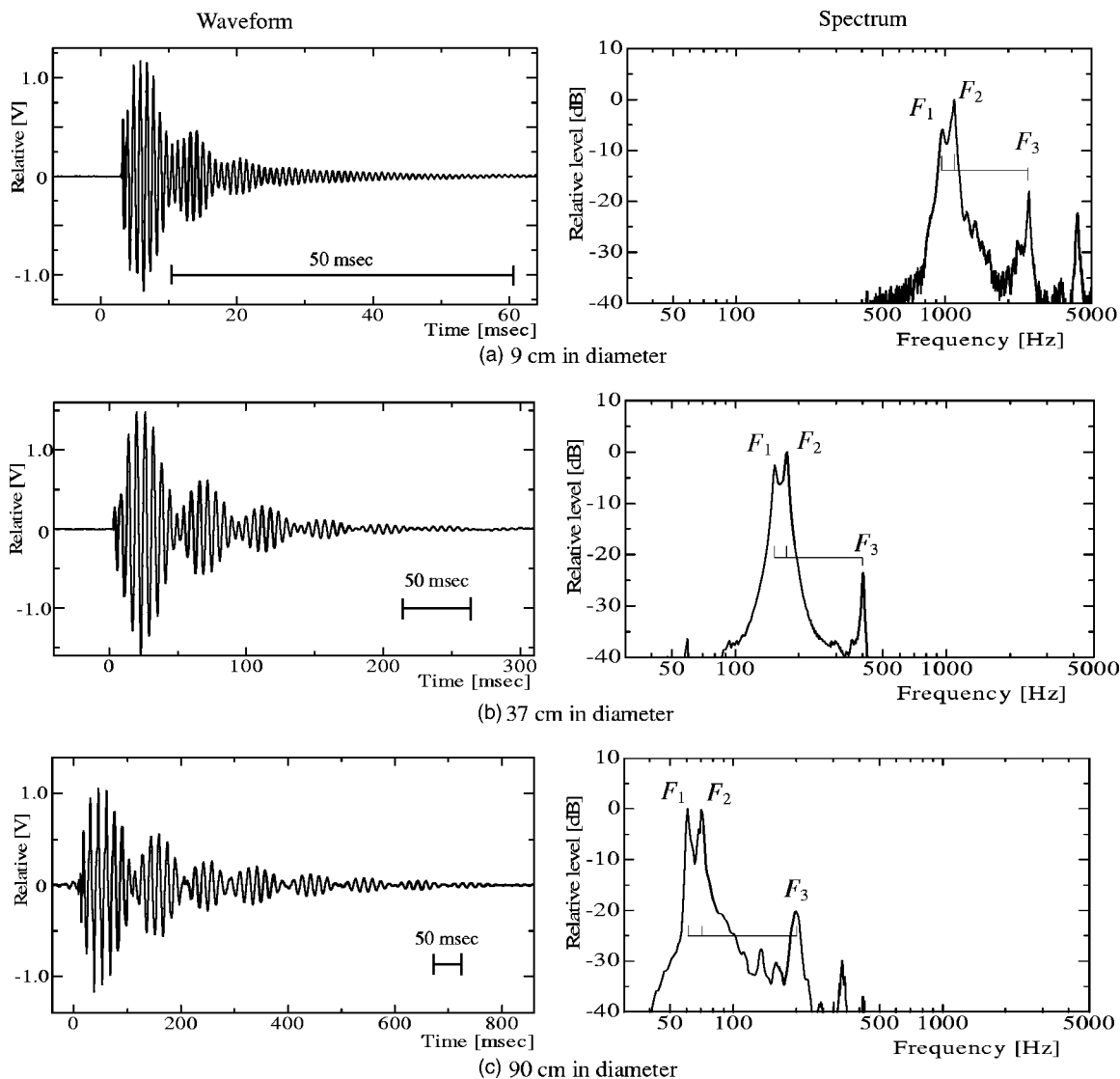


FIG. 2. Waveforms and spectra of three kinds of mokugyo: (a) 9 cm in diameter, (b) 37 cm, and (c) 90 cm.

mokugyo. Figure 3 shows a correlation diagram between F_1 and F_2 for these 174 mokugyo. F_1 and F_2 show a linear correlation, and its correlation coefficient is 0.997. The average ratio of F_1 and F_2 (i.e., the average of F_2/F_1) is ap-

proximately 1.15. These results suggest that the period of the beat observed in the waveform is $1000/(F_2 - F_1)$ (ms).

B. Vibration of the body

1. Experimental method

The mokugyo shown in Fig. 2(b) was used in this measurement. The vibration of the driving point of the body was measured using an acceleration pickup (RION: PV-90V) set 2 cm above the contact point for the drumstick along with the surface. The mokugyo was struck with a drumstick at the contact area shown in Fig. 1. The output acceleration signal was converted to a velocity signal via a charge amplifier (SHOWA: 4006).

2. Results

Figure 4 shows the waveform of the vibration velocity of the body, in which the rising edge of the waveform (i.e., when the drumstick contacts the body) is set to "0" in the time scale. A beat was observed in the waveform. The period

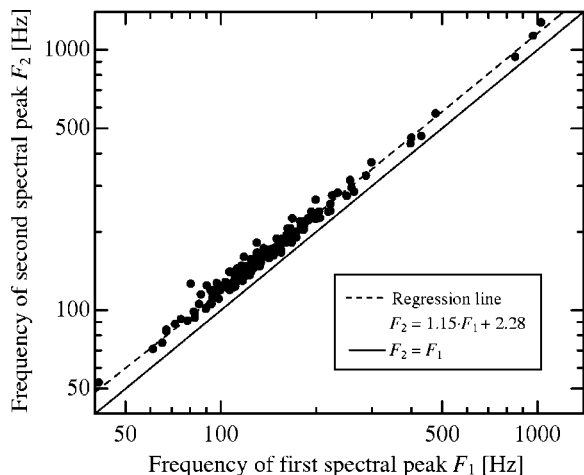


FIG. 3. Correlation between second peak F_2 and first peak F_1 of the sound spectra of 174 mokugyo.

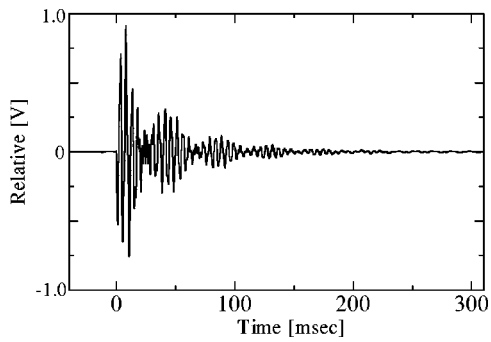


FIG. 4. Vibration velocity of the driving point of the body.

of the beat, 46 ms, is approximately equal to that of the sound wave in Fig. 2(b). As is seen in Fig. 2(b), the phase difference of the beat between the vibration and the sound is 26 ms, which is the sum of half the period of the beat and the propagation time to the observation point. Half the period was 23 ms, and the propagation time was 3 ms since the observation point for the sound was set at a distance of 100 cm away from the mokugyo. These results suggest that the energy of the vibration is exchanged alternately from the body to the Helmholtz resonator system. This energy exchange is caused by the coupled vibration of the two systems with the same eigenfrequency.

C. Effect of impact forces

1. Experimental method

Three kinds of drumsticks were used. Their tips were wrapped in thread, leather, and rubber. The length and mass of the drumstick with the thread tip were 23 cm and 26 g, those of the drumstick with the leather tip were 40 cm and 173 g, and those of the drumstick with the rubber tip were 55 cm and 250 g. The impact force was measured using an impedance head (RION: PF-60) set on a rigid wall. The impedance head was struck with the drumstick by an experimenter. The experimenter kept the driving force at 170 N through all measurements. The output signal was recorded via a charge amplifier (SHOWA: 4006).

Following the measurements of the impact forces, the sound spectra of a mokugyo with a 42-cm diameter were measured using two kinds of drumsticks: a leather-wrapped drumstick and a thread-wrapped drumstick. The method of recording and analyzing the measured sounds was the same as described in Sec. II A.

2. Results

Figure 5(a) shows the impact forces as a function of time. The force duration varies from 1 to 6 ms according to the material wrapped around the tip. Figure 5(b) shows the spectra of the impact forces. The frequency range varies according to the wrapping material.

Figure 6 shows the sound spectra of a mokugyo struck with two kinds of the drumsticks. In the frequency region less than $F_3 = 517$ Hz, the sound spectrum for the drumstick wrapped in thread corresponds to that for the one wrapped in leather. At higher frequencies the spectrum level for the

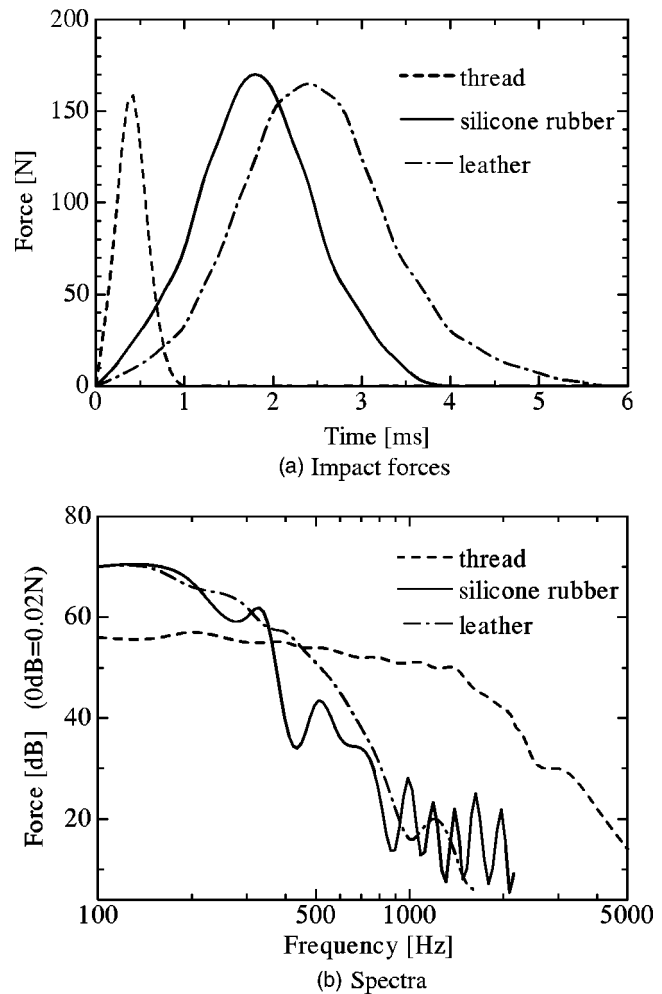


FIG. 5. Waveforms and spectra of impulsive forces from three kinds of drumsticks. (a) Impact forces. (b) Spectra.

drumstick wrapped in leather decreases. The result is consistent with the frequency range of the impact force.

D. Directional response

1. Experimental method

The mokugyo shown in Fig. 2(b) was used in this measurement, and its resonance frequencies are $F_1 = 155$ Hz,

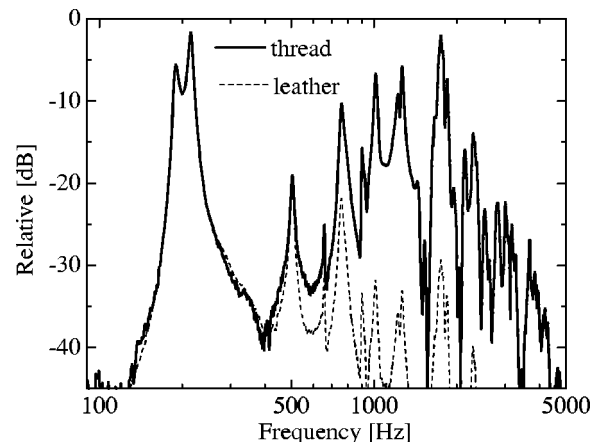


FIG. 6. Sound spectra from two kinds of drumsticks.

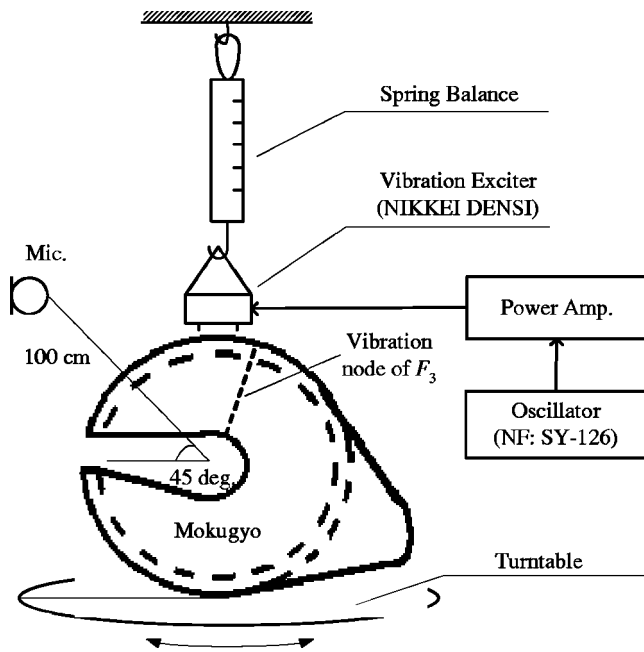


FIG. 7. Block diagram of directivity measurement setup.

$F_2 = 177$ Hz, and $F_3 = 402$ Hz. Figure 7 shows the block diagram of the measurement system. The mokugyo was put on a turntable and driven steadily with sinusoidal waves corresponding to each resonance frequency. The sound pressure level was measured as the turntable rotated. The output signal from the oscillator (NF: SY-126) was amplified and transmitted to a vibration exciter (Nikkei Denshi: Golden Tone) whose mass was compensated for by a spring balance. The body of the mokugyo was excited by the vibration exciter fixed on the contact area. A microphone was set at a distance of 100 cm away from the side port at a 45° angle from the horizontal plane. Also, an acceleration pickup (RION PV-90V) was used to investigate the vibration at various points on the surface of the body. This acceleration signal was recorded via a charge amplifier (SHOWA: 4006).

2. Results

Figure 8 shows the directivity for each resonance frequency. F_1 and F_2 are omnidirectional while F_3 shows a figure-8 pattern. This figure-8 pattern of F_3 was caused by the divided vibration of the body since the vibration node, which divides the upper part of the body in two, was observed as shown in Fig. 7.

E. Sound intensity measurement

1. Experimental method

The mokugyo shown in Fig. 2(b) was used. A two-dimensional sound intensity vector was measured by exciting the body using the same method described in Sec. IID. The measurement plane was a vertical one dividing the body in half. One hundred eighty-five measuring points, which were allocated on the measuring plane at 2-cm intervals around the side port, were used. The sound intensity vector was measured using the two rotating microphone method.^{9,10}

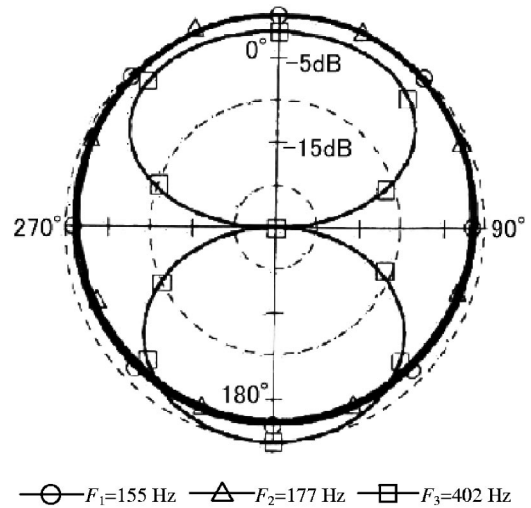


FIG. 8. The directional patterns of three resonant frequencies.

2. Results

Figure 9 shows the sound intensity vector around the side port at frequencies F_1 , F_2 , and F_3 . The arrows and their lengths indicate the directions and the levels (dB) of the vectors. The levels shown in Fig. 9(c) are compensated by +30 dB, since these levels are lower than the others.

For $F_1 = 155$ Hz and $F_2 = 177$ Hz, many of the vectors spread radially from the side port, and no vectors spread directly from the body. Therefore, the vibration mode at F_1 and F_2 would be a simple point sound source corresponding to the vibration of the air in the side port. This result is consistent with the directivity shown in Fig. 8. On the other hand, for $F_3 = 402$ Hz, the vectors from the side port are hardly observed, and most of the vectors are radiated from the body. This result suggests that the figure-8 pattern of F_3 shown in Fig. 8 is due to the divided vibration of the body.

III. THEORETICAL ANALYSIS

As shown in Sec. II, the sound of the mokugyo involves two close resonance frequencies, F_1 and F_2 . In this section, the relationship between these two resonance frequencies and the structure of the mokugyo is discussed based on a physical model. This structure composed of the body, cavity, and side port is analogous with that of the bass reflex loudspeaker,^{11,12} and a guitar,¹³⁻¹⁵ since there is coupling between vibrations of plates and the internal air cavity they enclose. An equivalent model of the mokugyo is proposed based on the theory of the bass reflex loudspeaker. Then, the mokugyo with a diameter of 37 cm shown in Fig. 2(b) is simulated acoustically.

A. Model of the mokugyo

Figure 10 shows the physical model of the mokugyo proposed in this paper, and Table I shows the parameters used in this model. The model is composed of two vibration systems: a mechanical vibration system driven by the drumstick and an acoustic vibration system composed of the cavity and the side port. In the mechanical vibration system, the effective vibration part of the body is assumed to be a circular vibratory board with equivalent mass m_m , which is fixed

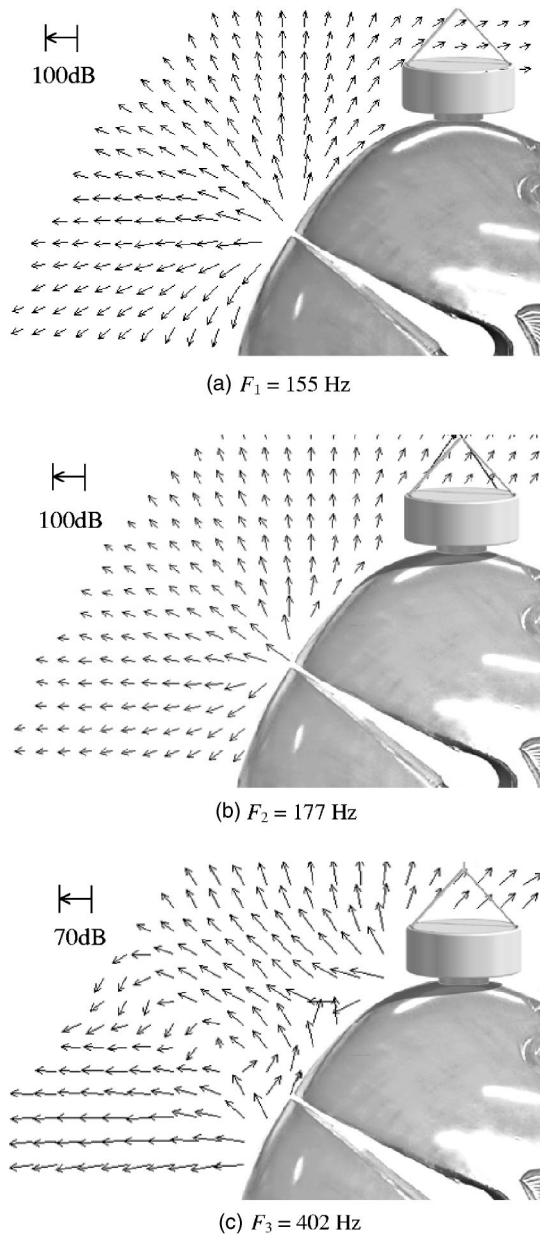


FIG. 9. Sound intensity vectors around the port: (a) $F_1 = 155$ Hz, (b) $F_2 = 177$ Hz, and (c) $F_3 = 402$ Hz.

on a rigid wall. m_m is derived from the sum of the mass of the board, m_d , and an additional mass radiated from the board, m_{ad} . The additional mass m_{ad} is a reaction force from the medium when the board is vibrating, and acts to increase the mass of the board. In this model, furthermore, the deflection forced by the drumstick is modeled using both a spring with equivalent stiffness, s_m , supporting the circular vibratory board and a mechanical resistance, r_m . The board is driven by an external force F due to the drumstick and is vibrated at a velocity v_d . In an acoustic vibration system, the cavity corresponds to the reciprocal of acoustic compliance, i.e., stiffness, s_c . The side port is modeled as a cylindrical hole hollowed out of a rigid wall. The hole is equivalent to the sum of the air mass, m_p , and an additional mass radiated from the side port, m_{ap} . The acoustic resistance of the side port is ignored in this model, since the resistance is much smaller than the mechanical resistance r_m .

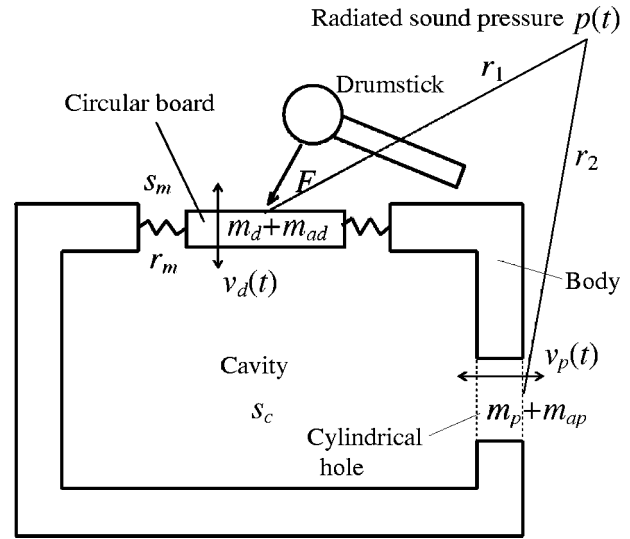


FIG. 10. Physical model of the mokugyo.

Figure 11 shows the electrical circuit equivalent to the model in Fig. 10. Coupling between the mechanical and acoustical systems is replaced by a transformer with a turn ratio equivalent to the area ratio of the effective vibration area of the body, S_d , and the cross-sectional area of the side port, S_p . Converting the secondary side (acoustical system) of the transformer into the primary side (mechanical system), the following simultaneous differential equations can be derived:

TABLE I. Parameter values used in the proposed model.

Variable	Symbol	Value
Mechanical system		
Effective vibration area of the board	S_d	0.07 m^2
Radius of the board	a_d	14.9 cm
Mass of the board	m_d	5.48 kg
Additional mass of the board	m_{ad}	0.616 kg
Equivalent mass	m_m	$m_d + m_{ad} = 6.10 \text{ kg}$
Equivalent stiffness	s_m	$6.76 \times 10^6 \text{ N/m}$
Equivalent resistance	r_m	$210 \text{ N} \cdot \text{s/m}$
Vibration velocity of the board	v_d	...
Acoustic system		
Volume of the port	W_p	$1.55 \times 10^{-3} \text{ m}^3$
Cross section area of the port	S_p	$1.50 \times 10^{-2} \text{ m}^2$
Radius of the port	a_p	6.9 cm
Length of the port	l_p	10.3 cm
Mass of the port's air	m_p	$1.87 \times 10^{-3} \text{ kg}$
Additional mass of the port	m_{ap}	$2.36 \times 10^{-3} \text{ kg}$
Equivalent mass converted into the mechanical system	m_1	$(S_d/S_p)^2(m_p + m_{ap}) = 0.092 \text{ kg}$
Volume of the cavity	W_c	$1.80 \times 10^{-2} \text{ m}^3$
Equivalent stiffness of the cavity	s_c	$1.0 \times 10^5 \text{ N/m}$
Particle velocity	v_p	...
Other parameters		
Impact force with the drumstick	F	170 N
Distance from the board to the observation point	r_1	130 cm
Distance from the side port to the observation point	r_2	100 cm
Air density	ρ_0	1.21 kg/m^3
Sound velocity	c	343 m/s

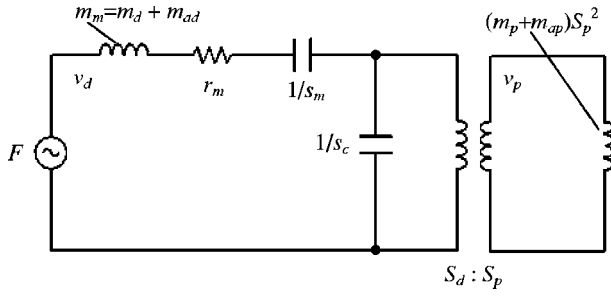


FIG. 11. Equivalent circuit of the physical model.

$$\begin{aligned}
 & m_m \frac{d}{dt} v_d(t) + r_m v_d(t) + s_m \int v_d(t) dt \\
 & + s_c \int \left\{ v_d(t) - \frac{S_p}{S_d} v_p(t) \right\} dt \\
 & = F(t),
 \end{aligned} \tag{1}$$

$$\hat{v}_d(\omega) = \frac{j\omega(s_c - m_l\omega^2)\hat{F}(\omega)}{m_m m_l \omega^4 - (s_c m_m + m_l s_c + m_l s_m)\omega^2 + s_c s_m - j(m_l r_m \omega^3 - s_c r_m \omega)} \tag{4}$$

and

$$\hat{v}_p(\omega) = \frac{j\omega s_c S_d \hat{F}(\omega) / S_p}{m_m m_l \omega^4 - (s_c m_m + m_l s_c + m_l s_m)\omega^2 + s_c s_m - j(m_l r_m \omega^3 - s_c r_m \omega)}, \tag{5}$$

where ω denotes the angular frequency, which is related to the vibration frequency f by $\omega = 2\pi f$. From the results shown in Sec. II E, the mechanical vibration system in the body and the acoustic center of the side port can be assumed to be simple point sound sources. Thus, the complex sound pressure $\hat{p}(\omega)$ at an arbitrary observation point is given by

$$\hat{p}(\omega) = \frac{j\rho_0\omega}{2\pi} \left\{ \frac{\epsilon^{-jkr_1}}{r_1} S_d \hat{v}_d(\omega) - \frac{\epsilon^{-jkr_2}}{r_2} S_p \hat{v}_p(\omega) \right\}, \tag{6}$$

where r_1 denotes the distance from the sound source of the mechanical vibration system, r_2 denotes that from the side port, ρ_0 denotes the density of air, and k denotes the wave constant given by $k = \omega c$. By applying Eqs. (4) and (5) to Eq. (6), the sound pressure at the observation point is given by

$$\hat{p}(\omega) = \frac{-\rho_0 S_d \omega^2 \hat{F}(\omega)}{2\pi} \left\{ \frac{(s_c - m_l\omega^2)\epsilon^{-jkr_1}/r_1 - s_c \epsilon^{-jkr_2}/r_2}{m_m m_l \omega^4 - (s_c m_m + m_l s_c + m_l s_m)\omega^2 + s_c s_m - j(m_l r_m \omega^3 - s_c r_m \omega)} \right\}. \tag{7}$$

B. Measurement of the model parameters

In this section, parameter values for the proposed model in Sec. III A are calculated using a mokugyo with a 37-cm diameter.

1. Model parameters

Table II shows the profile of the mokugyo used. Assuming the shape of the mokugyo is spherical, the surface area of the hemisphere (0.215 m^2) can be calculated from the diameter of the mokugyo. The area of the circular vibratory board, S_d , corresponding to the effective vibration area of the body in the model is supposed to be $\frac{1}{3}$ the surface area of the hemisphere. When $S_d = 0.07 \text{ m}^2$, the radius of the vibratory board, a_d , is 15.1 cm. In the proposed model, since the side

$$m_l \frac{S_p}{S_d} \frac{d}{dt} v_p(t) + s_c \int \left\{ \frac{S_p}{S_d} v_p(t) - v_d(t) \right\} dt = 0,$$

where $v_d(t)$ denotes the vibration velocity of the circular board, $v_p(t)$ denotes the particle velocity of the side port, and m_l denotes the equivalent mass of the side port converted to the mechanical system. m_l is given by

$$m_l = \left(\frac{S_d}{S_p} \right)^2 (m_p + m_{ap}). \tag{2}$$

The initial conditions of Eq. (1) are

$$v_d(0) = v_p(0) = \int_{-\infty}^0 v_d(0) dt = \int_{-\infty}^0 v_p(0) dt = 0. \tag{3}$$

From Eq. (1), the Fourier transforms of $v_d(t)$ and $v_p(t)$, that is, $\hat{v}_d(\omega)$ and $\hat{v}_p(\omega)$, are

port of the body is assumed to be a circular cylinder, its radius, a_p , is 6.9 cm and length, l_p , is 10.3 cm, where the cross section, S_p , of the side port (i.e., the cross section of the circular cylinder) is assumed to correspond to the actual area of the side port opening. The maximum force driven with the drumstick, F_{\max} , is set at 170 N as shown in Fig. 5.

TABLE II. Profile of the mokugyo used in the simulation.

Parameter	Actual measurement
Diameter of the mokugyo	37 cm
Aperture area of the side port	$1.50 \times 10^{-2} \text{ m}^2$
Capacity of the side port	$1.55 \times 10^{-3} \text{ m}^3$
Capacity of the cavity	$1.80 \times 10^{-2} \text{ m}^3$

The distance from each source to observation points r_1 and r_2 , the density of air, ρ_0 , and the sound velocity, c , are set at $r_1 = 130$ cm, $r_2 = 100$ cm, $\rho_0 = 1.21$ kg/m³, and $c = 343$ m/s.

2. Mechanical system parameters

a. *Equivalent mass of the circular vibratory board: m_d .* From the original first resonance frequency of the body, F_1 , and the resonance frequency, F'_1 , produced by fixing an arbitrary additional mass, m'_d , on the surface of the mokugyo, the equivalent mass of the vibratory board, m_d , is given by¹¹

$$m_d = m'_d \frac{F_1'^2}{F_1^2 - F_1'^2}. \quad (8)$$

$F_1 = 155$ Hz was used for the mokugyo as shown in Fig. 2(b). Here, F'_1 was measured and found to be 152.2 Hz, when the additional mass ($m'_d = 0.20$ kg) was experimentally fixed on the driving point of the body. Thus, $m_d = 5.48$ kg was obtained from Eq. (8).

b. *Additional mass of the board: m_{ad} .* The additional mass, m_{ad} , due to the radiation from the board, is derived from the following equation,¹²

$$m_{ad} = \frac{16}{3} \rho_0 a_d^3, \quad (9)$$

which expresses an inductive component of the radiated impedance produced by the circular vibratory board.¹² From Eq. (9), $m_{ad} = 0.3269$ kg.

c. *Equivalent stiffness: s_m .* From F_1 and m_d , the equivalent stiffness of the board s_m (Ref. 11) is

$$s_m = 4\pi^2 F_1^2 m_d. \quad (10)$$

From Eq. (10), $s_m = 6.76 \times 10^6$ N/m.

d. *Equivalent resistance: r_m .* The quality factor Q_1 of the first resonance frequency is 25.4, since Q_1 is $F_1 / \Delta F_Q$,¹¹ where ΔF_Q is the bandwidth when the sound pressure level becomes $1/\sqrt{2}$ times the sound pressure level at F_1 . Thus, the equivalent resistance of the mechanical system r_m (Ref. 11) is

$$r_m = \frac{2\pi F_1 m_d}{Q_1}. \quad (11)$$

From Eq. (11), $r_m = 210$ N·s/m.

3. Acoustical system parameters

a. *Stiffness of the cavity: s_c .* From the volume of the cavity, W_c , the area of the vibratory board, S_d , the sound velocity, c , and the density of air, ρ_0 , the stiffness of the cavity, s_c , is

$$s_c = \frac{\rho_0 c^2 S_d^2}{W_c}. \quad (12)$$

From Eq. (12), $s_c = 3.86 \times 10^4$ N/m.

b. *Air mass in the side port: m_p .* Multiplying the volume of the side port, W_p , by the density of air, ρ_0 , gives $m_p = 1.87 \times 10^{-3}$ kg.

c. *Additional mass of the side port: m_{ap} .* From Eq. (9), $m_{ap} = 2.13 \times 10^{-3}$ kg. This value was obtained in the same way as the additional mass of the vibratory board, m_{ad} .

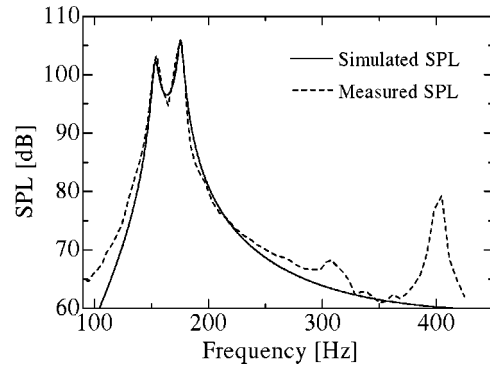


FIG. 12. Spectra of simulated and actual mokugyo.

C. Simulation

The sound spectrum of the mokugyo was simulated based on Eq. (7). Some of the parameter values, however, may be different from the actual values, since the shapes of the cavity and the side port assumed in the proposed model differ from the actual ones. W_c may also vary when struck by the drumstick. In this simulation, therefore, the additional mass of the vibratory board, m_{ad} , the stiffness of the cavity, s_c , and the additional mass of the side port, m_{ap} , were considered to be variables. These three parameters were optimized so as to match the simulated spectrum to the actual one.

First, the spectrum was calculated using Eq. (7) with the values of m_{ad} , s_c , and m_{ap} calculated in Sec. III B. Second, the optimal values for these three parameters were found by varying the values repeatedly. The convergence criteria were that the simulation error was within 1% for F_1 , F_2 , and the spectrum level difference at these frequencies. Here, $F_1 = 155$ Hz, $F_2 = 177$ Hz, and the level difference was 2.72 dB for the original sound spectrum as shown in Fig. 2(b).

Figure 12 shows the result obtained through the procedure described above. The final characteristics given by this simulation indicate that $F_1 = 155$ Hz, $F_2 = 177$ Hz, and the spectrum level difference is 2.72 dB. The result shows a correspondence between the simulated spectrum and the original one. The convergence values of the three parameters are $m_{ad} = 0.616$ kg, $s_c = 1.0 \times 10^5$ N/m, and $m_{ap} = 2.36 \times 10^{-3}$ kg.

Using these convergence values, the eigenfrequency of the mechanical system f_m is $1/2\pi\sqrt{s_m/m_m} = 167.5$ Hz, and that of the acoustic system f_c is $1/2\pi\sqrt{s_c/m_1} = 165.9$ Hz. These two frequencies are approximately equal. These results suggest that the mechanical vibration of the body and the acoustic vibration of the Helmholtz resonator are robustly coupled with each other, and that the beat of the waveforms shown in Fig. 2(b) is due to the mutual interaction of these two vibration systems.

From these results, the body, cavity, and side port create a structure for producing the two close resonance frequencies F_1 and F_2 . We assume that these two frequencies are the most important for the timbre of the mokugyo.

IV. COMPARISON OF THEORY AND EXPERIMENT

In the final tuning process of the mokugyo, an artisan chisels three areas called “innermost,” “side port,” and

TABLE III. Measured and estimated F_1 and F_2 before and after chiseling.

Part chiseled		“innermost”		“side port”		“front”	
Weight of sawdust		227.6 g		7.29 g		205.5 g	
		F_1 (Hz)	F_2 (Hz)	F_1 (Hz)	F_2 (Hz)	F_1 (Hz)	F_2 (Hz)
Measured	before chiseling	155.0	177.0	154.1	175.4	153.1	174.5
	after chiseling	154.1	175.4	153.1	174.5	153.4	175.1
Simulated	before chiseling	155.0	177.0	154.1	176.1	153.0	175.4
	after chiseling	154.1	176.1	153.0	175.4	153.4	176.0
Relative error		-0.03%	-0.38%	0.06%	-0.52%	0.05%	-0.52%
Main factors		$\Delta s_c = -2.08 \times 10^3$ N/m		$\Delta m_p = +1.49 \times 10^{-5}$ kg		$\Delta m_d = -6.90 \times 10^{-2}$ kg	
				$\Delta m_{ap} = +7.5 \times 10^{-5}$ kg			

“front” shown in Fig. 1. In this experiment, these areas were chiseled by an artisan and the sound spectra were measured before and after chiseling. The accuracy of the simulation with the proposed model described in Sec. III was examined using the results of the measurements during the actual chiseling.

A. Experiment and simulation conditions

1. Measurement experiment method

A mokugyo with a diameter of 37 cm was used for this experiment. The material of the body was camphorwood and its density was 0.16 g/cm³. The experimental procedure was as follows:

- The sound spectrum of the mokugyo was measured by the same method as described in Sec. II A.
- The “innermost” was chiseled by an artisan until he felt that the timbre changed.
- The mass of wood chips chiseled by the artisan and the sound spectrum for the chiseled mokugyo were measured.
- Procedures (a)–(c) were carried out for the “side port” and “front.”

2. Simulation method

The difference between the volume before and after chiseling the “innermost,” $\Delta W_{innermost}$, can be obtained from the mass of the wood chips and the density of the wood. $\Delta W_{innermost}$ is equal to the increase in volume of the cavity, ΔW_c , and the parameter corresponding to ΔW_c in the proposed model is the stiffness of the cavity s_c . Thus, substituting $\Delta W_{innermost}$ into ΔW_c in Eq. (12) gives the variation in stiffness, Δs_c . The simulated value of F_1 and F_2 can be obtained by substituting Δs_c into Eq. (7).

The decrease in volume of the “side port,” $\Delta W_{side\ port}$, can be also obtained from the mass of the wood chips and the density of the wood. $\Delta W_{side\ port}$ is equal to the increase in the volume of the side port, W_p . Thus, the variation of air mass in the side port, Δm_p , can be written as $\Delta W_{side\ port} \times \rho_0$. In addition, the additional mass of the side port, m_{ap} , increases since the cross section S_p of the side port increases due to the chiseling. Assuming that the length of the cylinder of the port l_p remains the same, the increase in the radius of the port, Δa_p , can be calculated, and then the additional mass, Δm_{ap} , can be obtained by using Eq. (9).

Considering the structure of the mokugyo, the mass of the “front” may vary with the decrease in the equivalent

mass, Δm_d , in the mechanical system. It would be difficult, however, to estimate a relationship between Δm_d and the mass of the wood chips from the “front.” Thus, in this simulation, Δm_d is assumed to be a variable parameter whose maximum value is equal to the mass of wood chips. The simulated values of F_1 and F_2 are calculated using Eq. (7). The optimal value of Δm_d is found using the same method as described in Sec. III C.

B. Results and discussion

Table III shows the weight of wood chips, measured and simulated resonance frequencies F_1 and F_2 before and after the chiseling, the relative error, and the main parameters thought to affect F_1 and F_2 . For “innermost” and “side port,” the measured F_1 and F_2 decreased after the chiseling, and the simulated values matched the measured ones well since the relative error was within 0.52%. The chiseling of the “side port” affects F_1 and F_2 significantly in spite of the fact that the mass of the chiseled wood chips, 7.6 g, is only 0.18% of the whole mass of the mokugyo, 4.3 kg. Such significant variations in F_1 and F_2 are due to the variation of the eigenfrequency of the Helmholtz system, which varies according to the variation of the volume of the “side port,” $\Delta W_{side\ port}$, and the mutual interaction of the vibration of the body and the Helmholtz system. In contrast, chiseling the “front” increased F_1 and F_2 , but the simulated values still matched the measured ones well since the relative error was within 0.52%. This seems to indicate that the top and bottom portions of the instrument are moving like the tines of a tuning fork or like portions of a slot drum. The vibrational mode of the wooden shell was not changed at its lowest sound peaks. The optimal value of Δm_d was found to be -6.90×10^{-2} kg, which is about $\frac{1}{3}$ of the mass of the wood chips. The reason why Δm_d becomes $\frac{1}{3}$ of the mass of the wood chips is a topic for further research. These results suggest that the proposed model can estimate ΔF_1 and ΔF_2 using the mass of the wood chips from the three areas selected here.

V. TIMBRE EVALUATION

As described in Sec. II A, beats are observed in the sound waveforms for all mokugyos, and their periods are $1000/(F_2 - F_1)$ (ms). In this section, the effects of F_1 and F_2 on the timbre of four kinds of mokugyos (with 9-, 18-, 37-, and 90-cm diameters) are investigated psychoacoustically with the semantic differential (SD) method. It is desir-

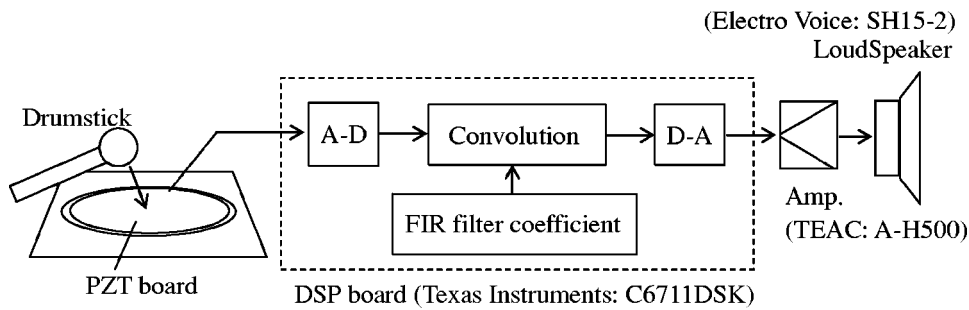


FIG. 13. Block diagram of the simulation synthesis method.

able for timbre evaluation that subjects can play the mokugyo. However, it is difficult to vary the resonance frequencies to arbitrary values by chiseling a real mokugyo. Thus, a system for simulating the sound of the mokugyo was made using a DSP board. This system can reproduce in real time simulated mokugyo sounds based on the model described in Sec. III.

A. Method of the simulation

Figure 13 shows a block diagram for synthesizing the sound using a DSP board (Texas Instruments: C6711DSK). The impact force of a drumstick was simulated as a voltage signal generated by striking a circular PZT board with a real drumstick. This signal was digitized with a sampling frequency of 2000 Hz and was transmitted to the DSP board. The parameter values of the proposed model for four kinds of real mokugyos were obtained by the same method as described in Secs. III B and C. Table IV shows sound characteristics of four real mokugyos and the values of F_1 and F_2 for each synthesized sound. Parameters in the proposed model were selected to match the F_1 and F_2 values shown in Table IV. The transfer function of the mokugyo body, i.e., FIR filter coefficients, was calculated for a sampling frequency of 2000 Hz. Then, the mokugyo sound was synthesized by convolution of the FIR coefficients with the impact force signal of the drumstick. This signal was reproduced by a loudspeaker (Electro Voice: SH15-2) via the D-A converter inside the DSP board.

B. Experiments

The timbre evaluation experiment based on the SD method was conducted in an anechoic room with a volume of 60 m³. Subjects reproduced each sound source with the system described in Sec. V A. The process was repeated until each subject felt that the evaluation was finished. To evaluate the four real mokugyos, the real mokugyos were also struck

TABLE IV. Sound characteristics of four real mokugyos and the values of F_1 and F_2 for the synthesized sounds.

Diameter of mokugyo (cm)	F_1 (Hz)	F_2 (Hz)	Center frequency F_c (Hz)	Values of $\Delta F = F_2 - F_1$ for synthesized sounds
9	984	1098	1041	0–200Hz(in50-Hzsteps)
18	392	427	409.5	0–60Hz(in10-Hzsteps)
37	155	177	166	0–40Hz(in10-Hzsteps)
90	70	80	75	0–20Hz(in 5-Hzsteps)

with a drumstick by the subject. The order of the experiments was randomized. The subjects were ten males and two females with normal hearing. In the experiment, ten rating words for a bipolar scale such as “bright” or “soft,” etc., and six rating words for a monopolar scale such as “like a wooden drum” or “comfortable,” etc., were used. The bipolar scale consists of nine points: -4 , Extremely; -3 , Very; -2 , “Dark” or “Hard,” etc.; -1 , A little; 0 , Midway; $+1$, A little; $+2$, “Bright” or “Soft,” etc.; $+3$, Very; $+4$, Extremely. The monopolar scale consists of seven points: 1 , Not at all; 2 , Not; 3 , Not too; 4 , A little; 5 , Good; 6 , Very; 7 , Extremely.

C. Results and discussion

Figure 14 shows the results of the timbre evaluation using the SD method for the real and the synthesized sounds. In this paper, only results for the mokugyo with a 37-cm diameter are shown. As shown in Figs. 14(a) and (b), the timbre of the real sound is mainly characterized by the rating words: “smooth,” “warm,” “gracious,” “rich,” “like a wooden drum,” and “comfortable.” The results for the synthesized sound with $\Delta F = 20$ Hz, which is closest to that of the real sound ($\Delta F = 22$ Hz), are similar to those for the real sound, that is, the timbre of the synthesized sound is close to that of the real sound. The six words mentioned above and “distinct” show higher scores. Here, $F_2/F_1 = 1.13$. On the other hand, scores for the words “distinct,” “clear,” “bright,” “cold,” and “unpleasant” are significantly higher for the synthesized sound with $\Delta F = 0$ Hz, and the shape of its polygon is significantly different from the real sound. The polygon of the sound for $\Delta F = 40$ Hz also differs from the others.

Figure 15 shows the relationship between F_2/F_1 and the average scores for the two monopolar rating words: (a) “like a wooden drum” and (b) “comfortable.” The results show that each evaluation score varies according to the value of F_2/F_1 . As shown in Fig. 15(a), the maximum scores for “like a wooden drum” are 3.25, 4.41, 4.42, and 3.29. Furthermore, the rates of change from the minimum score to the maximum one are 0.24, 0.54, 0.65, and 0.15 in order of increasing mokugyo size. Thus, the maximum score and the rate of change for 18- and 37-cm mokugyos are higher than those for 9- and 90-cm mokugyos. The values of F_2/F_1 for the maximum scores are 1.16, 1.08, 1.13, and 1.17. These values are close to the slope (1.15) of the regression line described in Fig. 3. The evaluation score of the real sound of the mokugyo with a 37-cm diameter is 4.25, whereas that of the synthesized sound is 4.42. As shown in Fig. 15(b), the

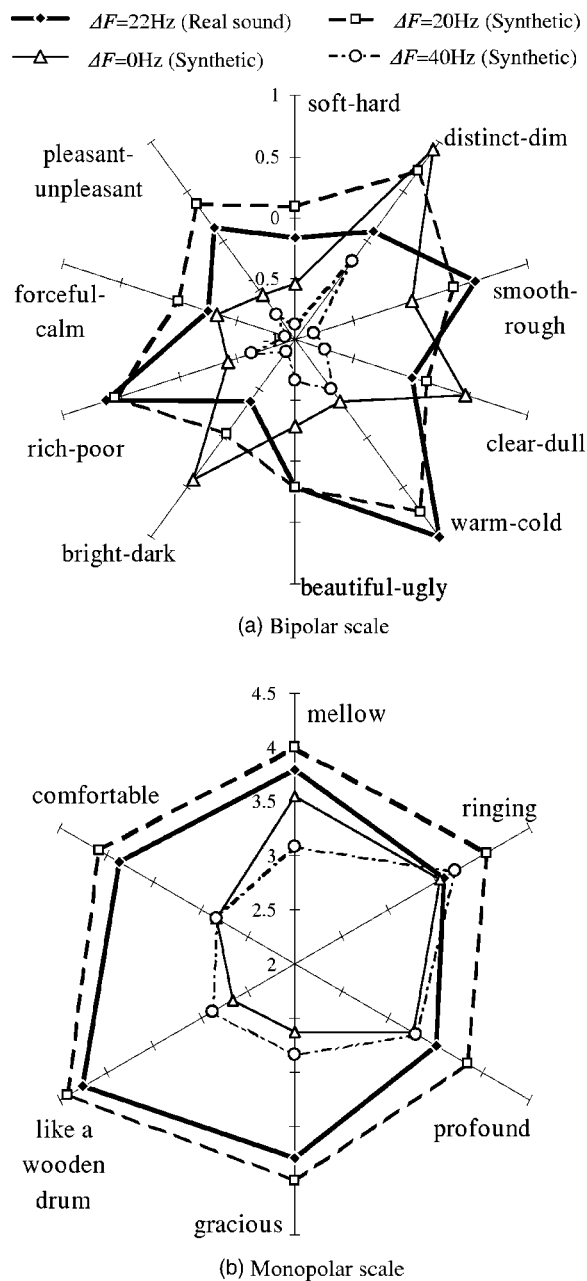


FIG. 14. Timbre evaluation results based on the SD method for a 37-cm mokugyo. (a) Results for a bipolar scale. (b) Results for a monopolar scale.

maximum scores for “comfortable” are 3.38, 4.27, 4.08, and 3.64 in order of increasing mokugyo size. The rates of change from minimum score to maximum one are 0.08, 0.40, 0.44, and 0.09 in order of increasing mokugyo size. Thus, the maximum score and the rate of change for 18- and 37-cm mokugyos are higher than those for 9- and 90-cm mokugyos. The values of F_2/F_1 for the maximum scores are 1.16, 1.08, 1.13, and 1.27. These values are also close to the slope 1.15. These results support our hypothesis that F_1 and F_2 are the most important factors affecting the timbre of a mokugyo. Also, the simulated sound based on the proposed physical model using a DSP board is effective.

VI. CONCLUSIONS

A traditional Japanese instrument, the mokugyo, which is made depending only on an artisan’s intuition, is investi-

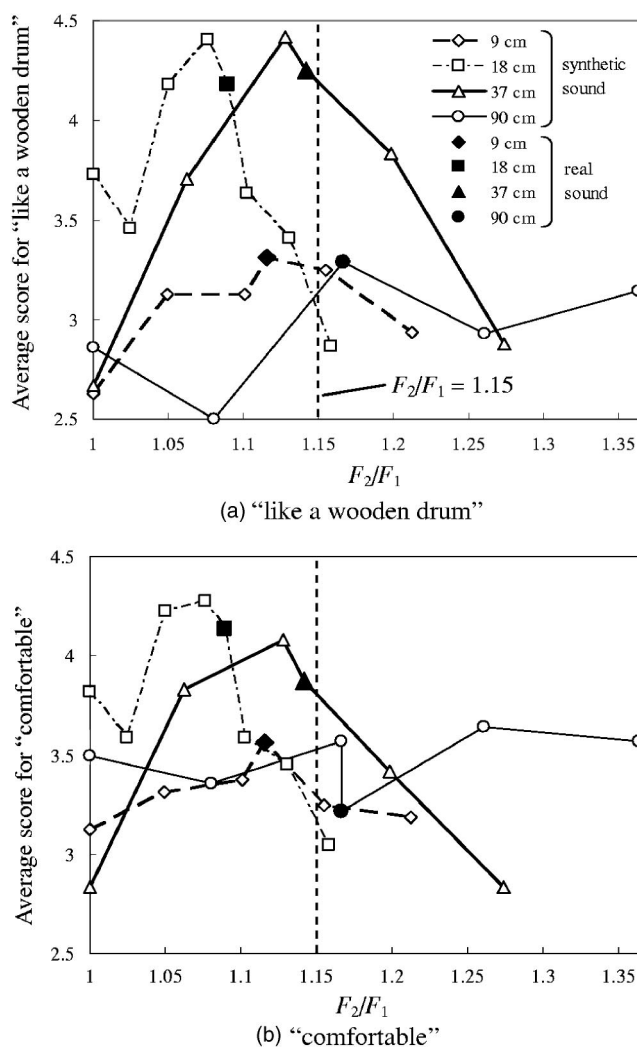


FIG. 15. Relationship between the evaluation score and F_2/F_1 : (a) “like a wooden drum” and (b) “comfortable.”

gated experimentally, theoretically, and psychoacoustically from an acoustic engineering standpoint. First, the physical characteristics of the mokugyo and the drumstick were experimentally measured and analyzed, and the following results were obtained:

- The sound of the mokugyo involves two close spectral peaks, F_1 and F_2 , and F_2/F_1 was approximately 1.15.
- The vibration of the driving point of the body shows a beat whose period is equal to the radiated sound wave.
- The duration of the impact force of the drumstick varies from 1 to 6 ms according to the material wrapped around the tip.
- The directivity of the F_1 and F_2 components is omnidirectional.
- The vibration mode of the F_1 and F_2 components is a point sound source.

Second, a physical model of the mokugyo was proposed. Because this model involves three variables which are difficult to measure, these variables are estimated using the real sound of the mokugyo. The proposed model was evaluated using a real mokugyo in the final tuning process. The results

showed that the model can estimate the difference between resonance frequencies before and after chiseling from the mass of wood chips from three areas: “innermost,” “side port,” and “front.”

Finally, the effects of F_1 and F_2 on the timbre of the mokugyo were investigated psychoacoustically using the semantic differential method. The results showed that when F_2/F_1 is approximately 1.15, the evaluation scores for the rating word, “like a wooden drum” or “comfortable,” showed higher scores.

It can be said that a mokugyo whose timbre is evaluated as a favorable sound can be made by controlling the physical characteristics corresponding to the timbre using the proposed model.

¹D. Adachi, K. Yamanaka, M. Sunohara, K. Furihata, T. Yanagisawa, and S. Yuasa, “Acoustic characteristics in a temple and timbre evaluation of virtual mokugyo” (in Japanese), *J. Inst. Electron., Inf. Commun. Eng.* **EA02-08**, 81–86 (2002).

²A. Chaigne and V. Doutaut, “Numerical simulations of xylophones. I. Time-domain modeling of the vibrating bars,” *J. Acoust. Soc. Am.* **101**, 539–557 (1997).

³V. Doutaut, D. Matignon, and A. Chaigne, “Numerical simulations of xylophones. II. Time-domain modeling of the resonator and of the radiated sound pressure,” *J. Acoust. Soc. Am.* **104**, 1633–1647 (1998).

⁴L. Rhaouti, A. Chaigne, and P. Joly, “Time-domain modeling and numerical simulation of a kettledrum,” *J. Acoust. Soc. Am.* **105**, 3545–3561 (1999).

⁵J. Obata and T. Tesima, “Experimental studies on the sound and vibration of drum,” *J. Acoust. Soc. Am.* **7**, 267–274 (1935).

⁶S. Ando and K. Yamaguchi, “Consideration on vibrations of Japanese traditional drums ‘Tsuzumi’” (in Japanese), *J. Acoust. Soc. Jpn.* **41(4)**, 239–247 (1985).

⁷S. Ando and K. Yamaguchi, “Acoustical studies of Japanese traditional drums ‘Tsuzumi’—Analysis, synthesis and subjective evaluation” (in Japanese), *J. Acoust. Soc. Jpn.* **41(6)**, 386–394 (1985).

⁸I. Aoki, S. Komatu, and M. Ise, “Effect of local loading on the acoustical property of Japanese bell” (in Japanese), *J. Acoust. Soc. Jpn.* **26(12)**, 562–571 (1970).

⁹T. Yanagisawa and N. Koike, “Cancellation of both phase mismatch and position errors with rotating microphones in sound intensity measurements,” *J. Sound Vib.* **113(1)**, 117–126 (1987).

¹⁰N. Koike and T. Yanagisawa, “Verification on possibility of separation of sound source direction by using a pair of pressure microphone” (in Japanese), *J. Inst. Electron., Inf. Commun. Eng.* **J77-A(1)**, 1–6 (1994).

¹¹H. Nakashima and T. Yamamoto, “Design procedures for the phase-inverted type loud speaker cabinets” (in Japanese), *Tech. J. Jpn. Broadcasting Corp.* **27**, 23–48 (1956).

¹²Y. Okada, S. Someda, and G. Hamano, “Analytical method of the phase-inverted type loud speaker” (in Japanese), *J. Inst. Electron., Inf. Commun. Eng.* **1**, 1–21 (1966).

¹³Ian M. Firth, “Physics of the guitar at the Helmholtz and first top plate resonances,” *J. Acoust. Soc. Am.* **61**, 588–593 (1977).

¹⁴G. Caldersmith, “Guitar as a reflex enclosure,” *J. Acoust. Soc. Am.* **63**, 1566–1575 (1978).

¹⁵O. Christensen, “Simple model for low-frequency guitar function,” *J. Acoust. Soc. Am.* **68**, 758–766 (1980).

Beating frequency and amplitude modulation of the piano tone due to coupling of tones

Bo Cartling^{a)}

Department of Physics, Royal Institute of Technology, SE-106 91 Stockholm, Sweden

(Received 2 June 2004; revised 28 December 2004; accepted 3 January 2005)

The influence on a piano tone from weak coexcitation of damped adjacent tones due to coupling via the bridge is studied. The frequency and amplitude modulation of the sound resulting from coexcitation of one strong and one or two weak tones is analyzed. One weak tone causes frequency and amplitude modulation of the sound, and two weak tones produce beating frequency and amplitude modulation, where the beatings of the two modulations are of opposite phase. By digital recording of the sound of piano tones, the appearance of these phenomena is verified. The audibility of the observed frequency and amplitude modulation is discussed in terms of previously determined detection thresholds. The beating character of both frequency and amplitude modulations, however, distinguishes the phenomena from those previously studied and prompts further psychoacoustic investigations. It is shown that detuning of unison strings may significantly increase the frequency deviation of the frequency modulation in conjunction with affected amplitude modulation. The modulatory effects of coupling to adjacent tones therefore may possibly be utilized in the tuning process. A coupling of tones analogous to the situation in a piano may arise in other stringed musical instruments transferring string vibrations to a soundboard via a bridge. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1861232]

PACS numbers: 43.75.Mn [NHF]

Pages: 2259–2267

I. INTRODUCTION

Several characteristics of the piano tone have been identified since the scientific study of the acoustics of the piano was initiated by von Helmholtz;¹ see Refs. 2–6 for reviews. The hammer being the distinguishing feature of the piano in relation to its predecessor, the harpsichord, the interaction between the hammer and the string received attention early¹ and has remained an active research area since.^{7–17} This interaction influences, e.g., the spectral and attack characteristics of piano tones. By designing a chromatic stroboscope,¹⁸ Railsback was able to determine the stretching of the equal-tempered scale, sharpening the treble and flattening the bass, consistently practiced by expert aural tuners.^{19,20} The physical rationale behind this was then established to be the inharmonicity of the string vibrations due to the stiffness of strings.^{21,22} The early and late phase of the decay of piano tones, named “prompt sound” and “aftersound,” respectively, were distinguished by Martin.²³ It was observed that expert aural tuners slightly detuned the strings of unisons.^{24,25} A seminal study of the physics of coupled strings by Weinreich^{26–28} revealed several important mechanisms. The prompt sound was ascribed to string vibrations of vertical polarization (in the configuration of a grand piano) with rapid energy transfer to the soundboard via the bridge, in particular symmetric motion of unison strings. The aftersound comes either from antisymmetric motion of unison strings, or from string vibrations of horizontal polarization only weakly coupled to the soundboard and feeding their energy back to the vertical mode of vibration. The aftersound can be affected by slight detunings of unison strings. As

regards coupling of unison strings via the bridge, Weinreich demonstrated that, for plane-polarized motion of two strings, a resistive coupling brings frequencies of the resulting normal modes closer together, whereas a reactive component of the coupling prevents a frequency crossing under adiabatic changes of system parameters. The degree of excitation of the different normal modes is determined by the initial conditions at the strike of the strings by the hammer. Weinreich also suggested that the objective for detuning the strings of unisons is to obtain a consistent aftersound for all tones by compensating for various irregularities in the sound generation by a piano. A recent study²⁹ considers the effect of localized anisotropy in the reactive part of the bridge admittance for a given partial frequency on so-called “false beating,” and suggests that the detuning of unison strings serves to reduce inheritance of false beats between partials of separate strings.

The bridge in a piano may transfer energy from excited string vibrations of a played tone to unstruck strings of adjacent tones. Such an indirect excitation will be weak compared to a direct blow by the hammer. Unless the sustaining pedal is engaged, the excitation will be further reduced by the dampers. The phase of the indirectly excited string vibrations starts out opposite to that of the directly excited string vibrations. The large frequency difference between the directly excited strings and the adjacent strings (5.9% for a semitone step), as compared with those of slightly detuned unison strings, makes the frequencies of the string vibrations less affected by the coupling. Nevertheless, the sound of a piano tone may be influenced by weak coexcitation of adjacent tones, and this is the subject of the present work. Already von Helmholtz¹ analyzed the amplitude variation when two sinusoids of different amplitude are superimposed

^{a)}Electronic mail: boc@kth.se

to produce beats, as well as the accompanying frequency modulation, as recently reviewed by Hartmann.³⁰

In this study, two cases of coexcited tones are first analyzed theoretically. In order to reach a high resolution in time, the analysis is based on the definition of frequency as the inverse of the time interval between successive zero crossings of the signal with the same change of sign. The frequency and amplitude modulation of the sound resulting from a strongly excited tone influenced by a weakly excited tone is first studied. Coexcitation of one strong tone and two weak tones is then shown to produce beating frequency and amplitude modulation of the resulting sound. By digital recording of piano tones, the appearance of these phenomena is verified.

II. COUPLING OF ADJACENT TONES

A. Coexcitation of one strong and one weak tone

Consider a coupling of two tones, such that a strong excitation of one tone causes a weak coexcitation of the other tone, initially at a phase opposite to that of the strongly excited tone. The two frequencies are assumed sufficiently far apart to remain nearly unaffected by the coupling. This situation may arise in a piano where a played tone may weakly coexcite a damped adjacent tone via the bridge. The resulting sound is described by

$$A(t) = \cos(\omega_0 t) + a \cos(\omega_1 t + \pi), \quad (1)$$

where ω_0 and ω_1 are the fundamental angular frequencies of the strong and weak tone, respectively. In a piano, the value of a is determined by both the weak excitation and the damping of the weak tone, making $a \ll 1$. Whereas decay rates of piano tones vary across the scale,²³ adjacent tones will have similar decay rates. Since these rates are much slower than that of the frequency and amplitude modulation to be analyzed here, the decay characteristics of the tones are not included. The implications of a time dependence of the relative amplitude a of the weak tone due to its damping will be discussed in Sec. III A. The frequency of the resulting sound $A(t)$ can be defined as the inverse of the time interval between successive zero crossings with the same change of sign. To determine this frequency, $A(t)$ is rewritten as

$$A(t) = \text{Re}(B(t)), \quad \text{where } B(t) = e^{i\omega_0 t} + a e^{i(\omega_1 t + \pi)}. \quad (2)$$

The summation of the two terms of $B(t)$ in a complex plane is illustrated in Fig. 1. A zero crossing of $A(t)$, from positive to negative sign, occurs at the positive imaginary axis. The presence of the weak tone of angular frequency ω_1 implies that successive zero crossings will occur at times deviating from those determined by only the strong tone of angular frequency ω_0 . Such a deviation is denoted by δt , and since $a \ll 1$, its magnitude satisfies $|\delta t| \ll 2\pi/\omega_0$, which will be utilized throughout the analysis below. The sign of δt depends on the relative phase of the two terms of $B(t)$, which is given by $\Delta\omega_0 t + \pi$, where $\Delta\omega_0 = \omega_1 - \omega_0$, as seen in Fig. 1. The fluctuations between advanced and retarded zero crossings, relative to the ones determined by the strong tone alone, thus give rise to a frequency modulation of the resulting sound $A(t)$ at an angular modulation frequency $|\Delta\omega_0|$.

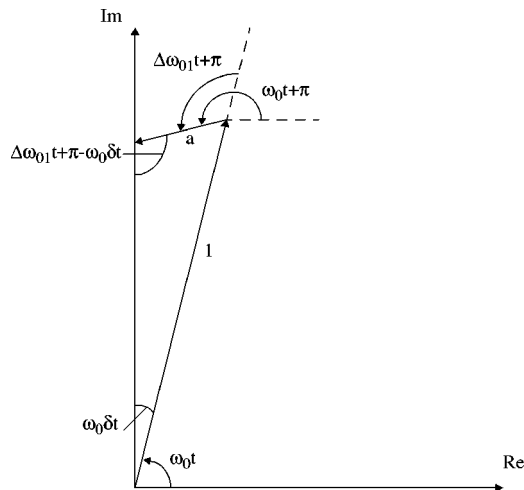


FIG. 1. An illustration of the superposition of a strong tone (vector of magnitude 1) at angular frequency ω_0 and a weak tone (vector of magnitude a , where $a \ll 1$) at an angular frequency $\Delta\omega_0$ from ω_0 , as described in a complex plane by Eq. (2). The depicted situation corresponds to a zero crossing of the resulting sound at a time δt in advance of a zero crossing of the strong tone alone. Further details are given in Sec. II A.

From Fig. 1, the following relation is obtained:

$$\frac{a}{\sin(\omega_0 \delta t)} = \frac{1}{\sin(\Delta\omega_0 t + \pi - \omega_0 \delta t)}, \quad (3)$$

which yields

$$\delta t = -\frac{1}{\omega_0} \text{atan} \left(\frac{\sin(\Delta\omega_0 t)}{\frac{1}{a} - \cos(\Delta\omega_0 t)} \right). \quad (4)$$

For $a \ll 1$, a simplified expression can be obtained as

$$\delta t \approx -\frac{1}{\omega_0} \text{atan}(a \sin(\Delta\omega_0 t)) \approx -\frac{a}{\omega_0} \sin(\Delta\omega_0 t). \quad (5)$$

To determine the time interval between two successive zero crossings, the following quantity is also defined:

$$\delta t' = -\frac{a}{\omega_0} \sin \left(\Delta\omega_0 \left(t + \frac{2\pi}{\omega_0} \right) \right). \quad (6)$$

The frequency of the resulting sound $A(t)$ can then be obtained as

$$\begin{aligned} \nu(t) &= \frac{1}{\frac{2\pi}{\omega_0} - (\delta t' - \delta t)} \\ &= \frac{1}{\frac{2\pi}{\omega_0} + \frac{2a}{\omega_0} \sin \left(\frac{\Delta\omega_0}{\omega_0} \pi \right) \cos \left(\Delta\omega_0 t + \frac{\Delta\omega_0}{\omega_0} \pi \right)}. \end{aligned} \quad (7)$$

This expression displays a periodic frequency modulation at angular frequency $|\Delta\omega_0|$ and shows that the modulation is not precisely sinusoidal. The expression interpolates between the zero crossings, that on average are separated by $2\pi/\omega_0$ s, corresponding to the crossing of lines through the origin of the complex plane at phase angles other than that of the imaginary axis. The frequency deviation (“amplitude”) $\Delta\nu$ of the frequency modulation, defined as the maximal differ-

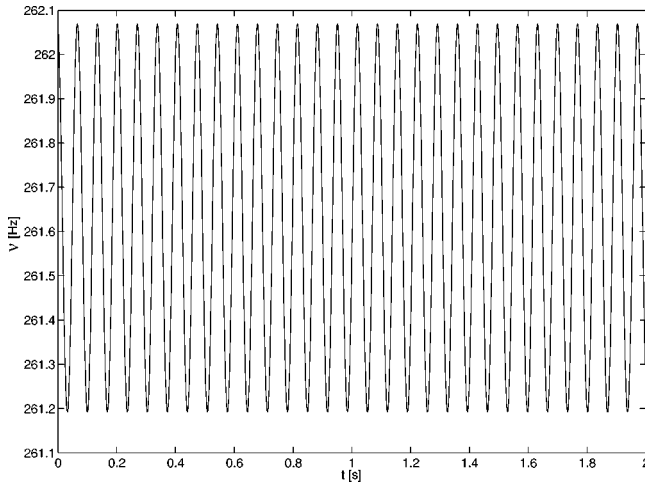


FIG. 2. Calculated frequency modulation of the fundamental of the C_4 tone due to weak coexcitation of the B_3 tone -30 dB relative to C_4 . The underlying expressions are derived in Sec. II A.

ence between the instantaneous frequency and the center frequency ν_0 , is obtained as

$$\Delta\nu = \frac{a\omega_0}{2\pi^2} \sin\left(\frac{|\Delta\omega_{01}|}{\omega_0}\pi\right). \quad (8)$$

An example is shown in Fig. 2, in which the strong tone is C_4 at $\omega_0/2\pi = 261.63$ Hz and the weak tone is B_3 at $\omega_1/2\pi = 246.94$ Hz. The value $a = 0.03$, corresponding to excitation of B_3 -30 dB relative to C_4 , is selected to be within the range of observed values, as discussed in Sec. III. The frequency modulation is at frequency $|\Delta\omega_{01}|/2\pi = 14.69$ Hz, and its frequency deviation is $\Delta\nu = 0.44$ Hz.

As evident, e.g., from Fig. 1, the resulting sound $A(t)$ will also be subject to an amplitude modulation with the same temporal characteristics as the frequency modulation, and can be expressed as

$$A(t) = \text{Re}(B(t)) \approx |B(t)| \cos(\omega_0 t - a \sin(\Delta\omega_{01} t)). \quad (9)$$

The amplitude modulation is described by the envelope function $|B(t)|$ of the faster temporal variation and can, for $a \ll 1$, be obtained as

$$|B(t)| \approx 1 - a \cos(\Delta\omega_{01} t). \quad (10)$$

The angular frequency of the amplitude modulation is thus given by $\Delta\omega_{01}$, and the maximal deviation of the amplitude from its average is equal to a . The ratio of the maximal amplitude deviation to the average amplitude is termed the modulation index of amplitude modulation. It is here given by a since the strong-signal amplitude is normalized to 1. The frequency and amplitude modulations are of same phase for $\Delta\omega_{01} > 0$ and of opposite phase for $\Delta\omega_{01} < 0$.

B. Coexcitation of one strong and two weak tones

Next, consider the coexcitation of one strong tone at angular frequency ω_0 and two weak tones at angular frequencies $\omega_1 < \omega_0$ and $\omega_2 > \omega_0$, under the same conditions as specified in Sec. II A. The resulting sound is described by

$$C(t) = \cos(\omega_0 t) + a \cos(\omega_1 t + \pi) + b \cos(\omega_2 t + \pi), \quad (11)$$

where $a, b \ll 1$. The two angular frequency differences are denoted $\Delta\omega_{01} = \omega_1 - \omega_0$ and $\Delta\omega_{02} = \omega_2 - \omega_0$. Based on the analysis in Sec. II A, the following quantities are defined:

$$\delta t_{01} = -\frac{a}{\omega_0} \sin(\Delta\omega_{01} t), \quad (12)$$

$$\delta t_{02} = -\frac{b}{\omega_0} \sin(\Delta\omega_{02} t), \quad (13)$$

$$\delta t'_{01} = -\frac{a}{\omega_0} \sin\left(\Delta\omega_{01}\left(t + \frac{2\pi}{\omega_0}\right)\right), \quad (14)$$

$$\delta t'_{02} = -\frac{b}{\omega_0} \sin\left(\Delta\omega_{02}\left(t + \frac{2\pi}{\omega_0}\right)\right). \quad (15)$$

The frequency of the resulting sound $C(t)$ can then be obtained as

$$\nu(t) = \frac{1}{\frac{2\pi}{\omega_0} - (\delta t'_{01} - \delta t_{01} + \delta t'_{02} - \delta t_{02})}. \quad (16)$$

For simplicity, consider $b = a$. It then follows that

$$\begin{aligned} \delta t_{01} + \delta t_{02} &= \frac{2a}{\omega_0} \cos\left(\frac{|\Delta\omega_{01}| + |\Delta\omega_{02}|}{2} t\right) \\ &\quad \times \sin\left(\frac{|\Delta\omega_{01}| - |\Delta\omega_{02}|}{2} t\right), \end{aligned} \quad (17)$$

and

$$\begin{aligned} \delta t'_{01} + \delta t'_{02} &= \frac{2a}{\omega_0} \cos\left(\frac{|\Delta\omega_{01}| + |\Delta\omega_{02}|}{2} \left(t + \frac{2\pi}{\omega_0}\right)\right) \\ &\quad \times \sin\left(\frac{|\Delta\omega_{01}| - |\Delta\omega_{02}|}{2} \left(t + \frac{2\pi}{\omega_0}\right)\right). \end{aligned} \quad (18)$$

Both Eqs. (17) and (18) include a faster variation at an angular frequency $(|\Delta\omega_{01}| + |\Delta\omega_{02}|)/2$, and a slower variation at an angular frequency $||\Delta\omega_{01}| - |\Delta\omega_{02}||/2$. The frequency of the resulting sound $C(t)$ is consequently periodically modulated at the angular frequency $(|\Delta\omega_{01}| + |\Delta\omega_{02}|)/2$, i.e., the average of the two angular frequency differences. The frequency change of the frequency modulation varies at an angular frequency $||\Delta\omega_{01}| - |\Delta\omega_{02}||$, i.e., the difference between the two angular frequency differences. A phenomenon resembling “beating” (which normally refers to the amplitude modulation caused by partials close in frequency) thus appears in the frequency modulation in this case. For a tone in the equal-tempered scale, the angular beat frequency is 0.33% of ω_0 , assuming that the weak tones are adjacent semitones. An example is shown in Fig. 3, in which the strong tone is C_4 at $\omega_0/2\pi = 261.63$ Hz and one weak tone is B_3 at $\omega_1/2\pi = 246.94$ Hz and the other weak tone is $C_4^\#$ at $\omega_2/2\pi = 277.18$ Hz. For the same reason as in Sec. II A, the values $a = b = 0.03$ are selected, corresponding to excitation of both B_3 and $C_4^\#$ -30 dB relative to C_4 . The frequency differences are given by $|\Delta\omega_{01}|/2\pi = 14.69$ Hz and $|\Delta\omega_{02}|/2\pi = 15.55$ Hz, so that the frequency modulation is at

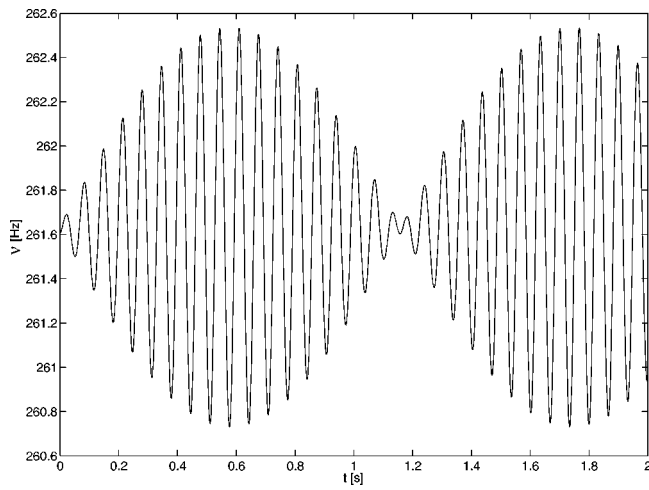


FIG. 3. Calculated frequency modulation of the fundamental of the C_4 tone due to weak coexcitation of the B_3 and $C_4^\#$ tones -30 dB relative to C_4 . The underlying expressions are derived in Sec. II B.

15.12 Hz and the beat frequency of the frequency modulation is 0.86 Hz. The frequency deviation approximately doubles (for $|\Delta\omega_{01}| \approx |\Delta\omega_{02}|$) compared with the case in Sec. II A.

As for the case of one strong and one weak tone described in Sec. II A, there is also an amplitude modulation of the resulting sound $C(t)$. Denoting the complex quantity corresponding to $C(t)$ by $D(t)$ [in analogy with Eq. (2)] and considering $b = a$, $C(t)$ can be expressed as

$$C(t) = \text{Re}(D(t)) = |D(t)| \cos(\omega_0 t - a \sin(\Delta\omega_{01} t) - a \sin(\Delta\omega_{02} t)). \quad (19)$$

The envelope function $|D(t)|$ describing the amplitude modulation can, for $a \ll 1$, be obtained as

$$|D(t)| \approx 1 - 2a \cos\left(\frac{|\Delta\omega_{01}| + |\Delta\omega_{02}|}{2} t\right) \times \cos\left(\frac{|\Delta\omega_{01}| - |\Delta\omega_{02}|}{2} t\right). \quad (20)$$

The amplitude modulation thus displays a faster variation at an angular frequency $(|\Delta\omega_{01}| + |\Delta\omega_{02}|)/2$ and a slower variation at an angular frequency $||\Delta\omega_{01}| - |\Delta\omega_{02}||/2$, the same two frequencies characterizing the frequency modulation. As for the frequency modulation, the slower variation can be referred to as a beating of the amplitude modulation at an angular frequency $||\Delta\omega_{01}| - |\Delta\omega_{02}||$. The modulation index of the amplitude modulation is given by $2a$, i.e., twice the value for the case in Sec. II A. Defining the relative phase shift between the frequency and amplitude modulations to be positive when the frequency modulation precedes the amplitude modulation, then the relative phaseshift switches from $+\pi/2$ to $-\pi/2$ at the maxima of the frequency-beat envelope and vice versa at the minima. The beatings of the two modulations are of opposite phase. An example of the amplitude modulation is shown in Fig. 4 for the same case as defined for Fig. 3. To fully describe the temporal development of piano tones, their decay characteristics should also be included.

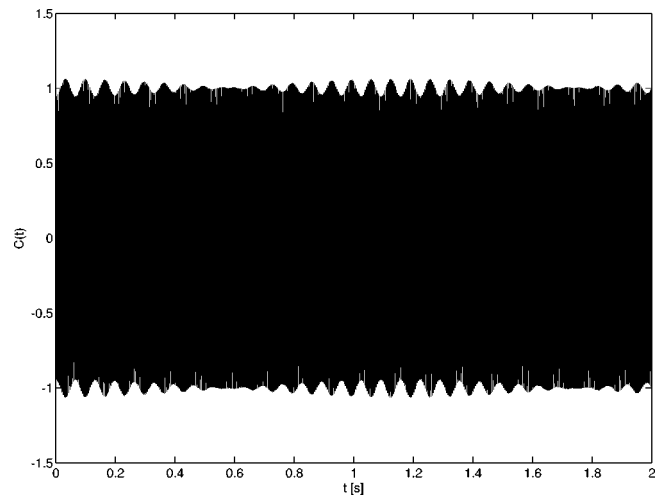


FIG. 4. Calculated amplitude modulation for the same case as in Fig. 3. The underlying expressions are derived in Sec. II B. Decay characteristics of the tones are not included.

III. PIANO TONES

A. Single-string tones

The situations treated in Sec. II, giving rise to frequency and amplitude modulation or beating frequency and amplitude modulation, may apply to piano tones. In the piano, tones other than the played ones are intended to be damped, unless the sustaining pedal is engaged. However, the frequency and amplitude modulation of the resulting sound is quite sensitive to the influence of even weakly excited tones, as demonstrated in Sec. II. A coupling of adjacent tones via the bridge in a piano can generate weakly excited tones along with a strongly excited tone. To investigate this possibility, measurements were performed by digital recording of the sound from an upright piano (Yamaha model M1E) via a microphone. The sampling rate of the recording was 48 kHz and the resolution 16 bits. As mentioned in the Introduction, it is well known that the vibration of a piano string is not perfectly periodic due to inharmonicity, which makes the frequencies of overtones slightly different from precise multiples of the fundamental frequency. To remove contributions to nonperiodicity from overtones, the recorded signal was bandpass filtered around the fundamental frequency of the played tone. This was accomplished by a fast Fourier transform (FFT) followed by an inverse FFT after retaining Fourier components across the desired range. To improve the spectral estimate, a Hann window was applied to the recorded signal before the FFT.³¹ The bandpass filter was realized by a square window of full width $0.2\nu_0$ around ν_0 , where ν_0 denotes the fundamental frequency of the played tone. This width allows the fundamentals of adjacent tones, a semitone step down or up, to contribute while suppressing overtones of the played tone as well as fundamentals of next-nearest semitones. The processed signal was then frequency analyzed in terms of the inverse time interval between successive zero crossings with the same change of sign. The frequency resolution is improved by linear interpolation of the zero crossings and can be estimated at 0.1 Hz for a C_4 tone at about 262 Hz.

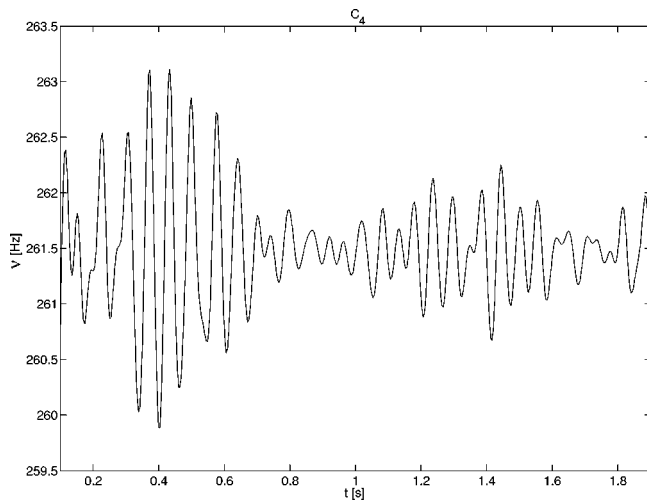


FIG. 5. Observed frequency modulation of the fundamental of the C_4 tone recorded from an upright piano. Each of the C_4 , B_3 , and C_4^\sharp tones is generated by a single string. The recorded sound is bandpass filtered around the fundamental of the C_4 tone, and the frequency is determined from successive zero crossings of the signal with the same change of sign. The time interval is 0.1 to 1.9 s from excitation. Further details of the recording and analysis are given in Sec. III A.

Figure 5 shows the result obtained by means of this analysis of the recorded sound of a C_4 tone. The adjacent tones, a semitone step down and up, are B_3 and C_4^\sharp , respectively. In this measurement, two of the three strings of each unison are damped, so that each tone is generated by a single string. To compensate for inharmonicity (“stretched octaves”), as described in the Introduction, the piano was tuned so that the equal-tempered frequencies of these three tones were reduced by 1 cent, i.e., C_4 at $\omega_0/2\pi = 261.47$ Hz, B_3 at $\omega_1/2\pi = 246.80$ Hz, and C_4^\sharp at $\omega_2/2\pi = 277.02$ Hz. The tuning was accomplished by means of a 1-s digital recording of the sound followed by a calculation of a 2001-point spectrum across a frequency interval of 20 Hz around the fundamental frequency by means of a chirp z-transform.³¹ The resolution of the spectrum is thus 0.01 Hz, so that the tuning accuracy is limited by the precision by which the tuning pins can be set, about ± 0.05 Hz. The two frequency differences are $|\Delta\omega_{01}|/2\pi = 14.67$ Hz and $|\Delta\omega_{02}|/2\pi = 15.55$ Hz, so that their average is 15.11 Hz and their difference 0.88 Hz. The figure displays a time interval from 0.1 to 1.9 s of a 2-s recording excluding the initial and final 0.1 s where the signal is strongly reduced by the Hann window applied prior to the FFT, as described above. During the 2-s period, the sound is of the prompt type before the aftersound appears. For a single string, it is vibrations of horizontal polarization that contribute to the aftersound. Figure 5 demonstrates a beating frequency modulation in fair agreement with the predictions from the analysis in Sec. II B of one strongly excited tone and two weakly excited tones. The observed maximal peak-to-peak frequency swing of about 3 Hz corresponds to a frequency deviation of about 1.5 Hz and a value of the a amplitude of about 0.05. (Recall that Figs. 3 and 4 refer to $a = 0.03$.) A time-dependent frequency deviation can be defined as the maximal frequency change from the center frequency ν_0 determined over a modulation period. A decrease of frequency deviation with time may derive from the damp-

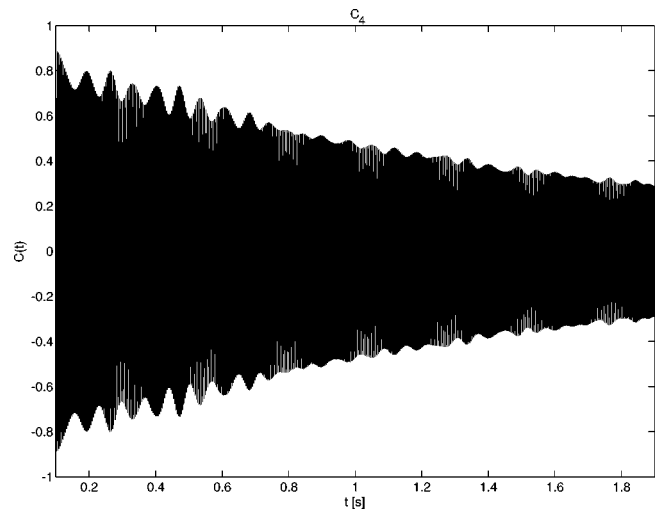


FIG. 6. Observed amplitude modulation of the fundamental of the C_4 tone for the same sound, and over the same time interval, as in Fig. 5. The recorded sound is bandpass filtered around the fundamental of the C_4 tone. Further details of the recording and analysis are given in Sec. III A.

ing of the weakly excited tones, i.e., the amplitude a may be a decreasing function of time. To display the amplitude modulation of the sound, the recorded signal was bandpass filtered as above but without the application of a Hann window prior to the FFT. The resulting signal is shown in Fig. 6 over the same time interval as in Fig. 5. Also the amplitude modulation is in fair agreement with the predictions in Sec. II B. Possible sources of deviations from the calculated behavior include a difference of the weak amplitudes a and b in Eq. (11). If one of these amplitudes dominates over the other, then the situation described in Sec. II A is more applicable. Additional time dependence of a and b , apart from their decay due to damping, may also interfere with the beating patterns. Noise affecting amplitude relationships and phase relationships between the component signals will cause irregularities in the frequency and amplitude modulation, in particular the beating patterns. Since the beating frequency is a small difference between larger frequency differences, a small relative change of these larger frequency differences implies a large relative influence on the beating frequency. The estimated amplitude value of the weakly excited adjacent tones may seem large. The excitation of the adjacent tones is, however, presumably linked to the total sound pressure of the played tone. Relating this excitation to the excitation of the fundamental of the played tone results in a larger ratio than that between the total sound pressures of adjacent and played tones.

An influence of the damping of the weak tones on the frequency deviation of the frequency modulation was also confirmed. The frequency deviation is reduced if further damping is imposed on the strings of the adjacent tones, and removal of damping increases the frequency deviation. Figure 7 displays the frequency modulation of the C_4 fundamental when the dampers are lifted off the strings of the adjacent B_3 and C_4^\sharp tones. The maximal peak-to-peak frequency swing increases to about 8 Hz.

One more example is shown in Fig. 8 which refers to the A_4 tone, with adjacent tones one semitone step away G_4^\sharp and

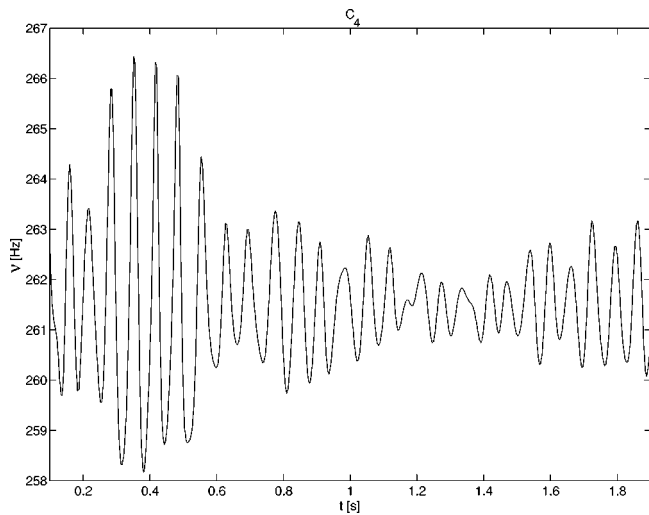


FIG. 7. Same as Fig. 5, but with the dampers lifted off the strings of the B_3 and $C_4^\#$ tones.

$A_4^\#$. As in the previous case, the measurement is performed with each tone generated by a single string. In this range, the piano is tuned according to equal temperament, so that A_4 is at $\omega_0/2\pi = 440.00$ Hz, $G_4^\#$ at $\omega_1/2\pi = 415.31$ Hz, and $A_4^\#$ at $\omega_2/2\pi = 466.16$ Hz. This yields the two frequency differences $|\Delta\omega_{01}|/2\pi = 24.69$ Hz and $|\Delta\omega_{02}|/2\pi = 26.16$ Hz, with an average 25.43 Hz and a difference 1.47 Hz. To ensure the prompt type of sound, the recording is restricted to 1 s. The displayed time interval omits the initial and final 0.1 s, for the same reason as in the previous example. The agreement with predictions is again fair. The observed maximal peak-to-peak frequency swing of about 5 Hz corresponds to a value of the a amplitude of about 0.05. The tones F_3 , F_4 , and C_5 were also measured to study average frequency differences of approximately 10, 20, and 30 Hz, in addition to the approximately 15 and 25 Hz for C_4 and A_4 , and an analogous degree of agreement with predictions was observed. The results obtained therefore indicate that weak co-

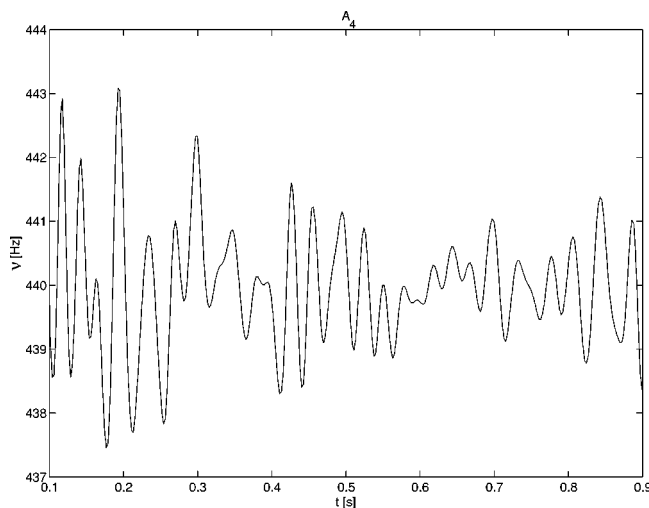


FIG. 8. Same as Fig. 5, but for the A_4 tone and the time interval 0.1 to 0.9 s from excitation. Each of the A_4 , $G_4^\#$, and $A_4^\#$ tones is generated by a single string.

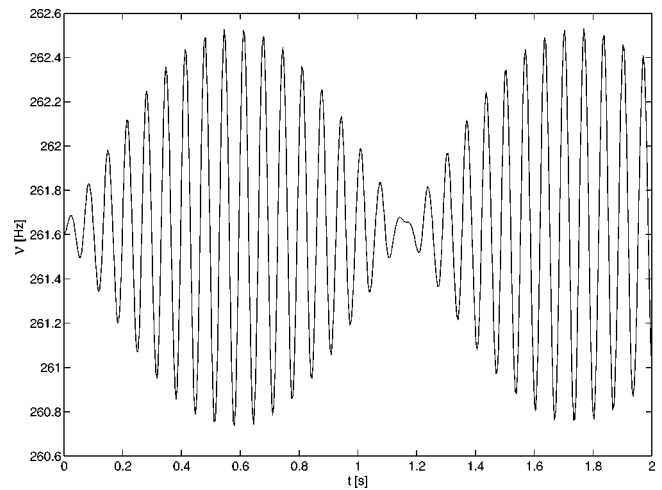


FIG. 9. Simulated frequency modulation of the fundamental of the C_4 tone due to weak coexcitation of the B_3 and $C_4^\#$ tones -30 dB relative to C_4 . Each of the tones is generated by three strings detuned -1 , 0 , and 1 cent from the respective fundamental frequency. Further details of the simulation are given in Sec. III B.

excitation of adjacent tones can contribute beating frequency and amplitude modulation to the piano tone.

B. Triple-string tones

All of the tones studied are generated by three unison strings, so that a slight detuning of these, as discussed in the Introduction, may produce up to 18, rather than two, frequency differences between the played tone and adjacent tones. Only nine of these frequency differences are independent, corresponding to the number of strings. To determine the effects of such a multitude of frequency differences on the frequency and amplitude modulation, computer simulations as well as measurements were performed. Figure 9 shows the frequency modulation of the fundamental of the C_4 tone as obtained by a simulation in which each of the three tones involved is generated by the three strings of a unison disregarding coupling effects within a unison. Fundamental frequencies, and the amplitudes of the weak tones, are as described in Sec. II B for Fig. 3. Within each unison, the three strings are detuned -1 , 0 , and 1 cent, respectively, from the fundamental frequency. Such a detuning pattern within the unisons does not cause a significant difference of the resulting frequency modulation from what is obtained for single-string tones, as seen by a comparison of Figs. 3 and 9.

Another detuning pattern of the unisons can, however, produce rather different frequency modulation. Figure 10 shows the result of a simulation of the C_4 tone, where the three strings of the C_4 unison are detuned -1 , 0 , and 2 cents from the fundamental frequency and the strings of the other two unisons involved are detuned -1 , 0 , and 1 cent from the respective fundamental frequency. With this type of detuning within the unisons, the amplitude modulation will be more affected as shown in Fig. 11 for the same case. For detuning patterns that produce zero transitions of the amplitude, rather large frequency deviations may occur at these. However, they will probably not be perceived due to the low amplitude. An example of observations near the latter situation is

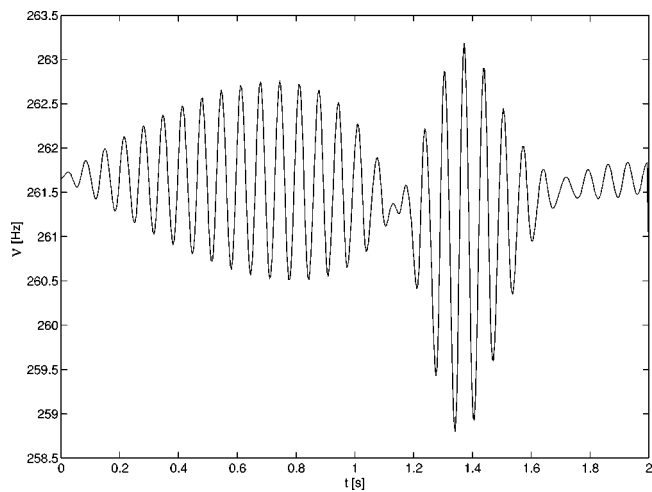


FIG. 10. Same as Fig. 9, but with the three strings of C_4 detuned -1 , 0 , and 2 cents from the fundamental frequency. The three strings of each of the B_3 and $C_4^\#$ tones are detuned -1 , 0 , and 1 cent from the respective fundamental frequency.

shown in Figs. 12 and 13, displaying frequency and amplitude modulation, respectively. Just after the recording of this signal, the nine fundamental frequencies involved were determined by the procedure described in Sec. III A for tuning and found to be 261.16 , 261.48 , and 261.51 Hz for the three unison strings of C_4 , 246.63 , 246.78 , and 246.92 Hz for B_3 and 276.88 , 277.09 , and 277.39 Hz for $C_4^\#$. As described in the Introduction, the frequencies of the normal modes that result from bridge coupling of strings in a unison may be different from the values to which the strings are tuned when sounding separately, and the excitation of the normal modes is determined by the initial conditions at the strike by the hammer.^{26–28}

IV. DISCUSSION

Frequency and amplitude modulations are phenomena commonly associated with sinusoidal modulation,³⁰ which produces Fourier components at frequencies deviating from the carrier frequency by positive and negative multiples of

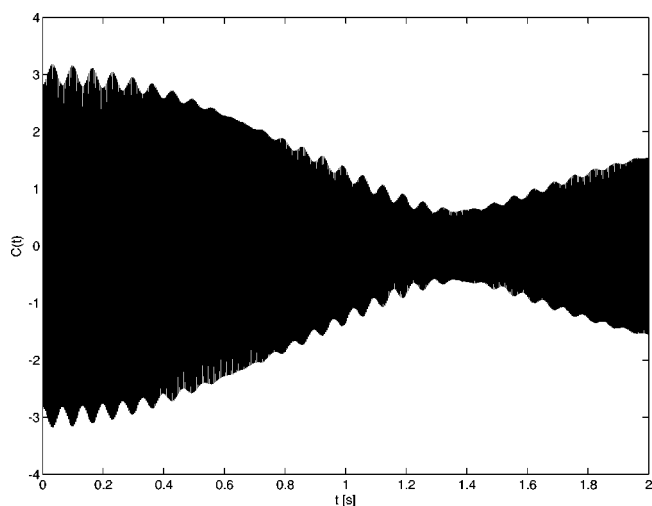


FIG. 11. Simulated amplitude modulation for the same case as in Fig. 10. Decay characteristics of the tones are not included.

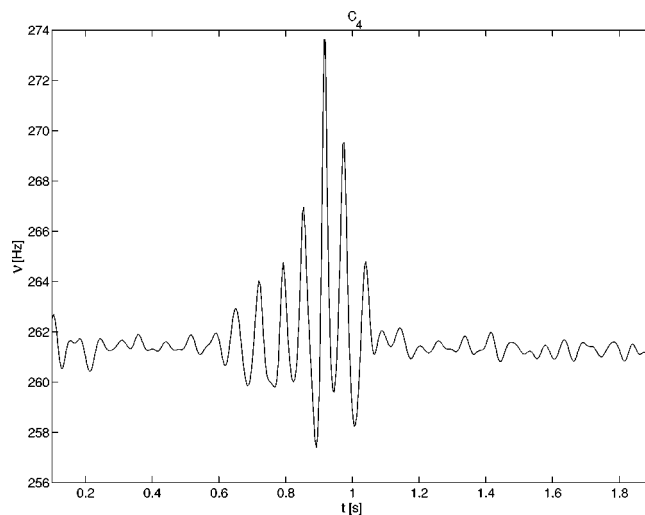


FIG. 12. Same as Fig. 5, but with each tone generated by three strings. The figure illustrates an example of detuning pattern generating rather large frequency deviations. The particular frequency values, to which the strings are tuned when sounding separately, are given in Sec. III B.

the modulation frequency (frequency modulation) or plus and minus the modulation frequency (amplitude modulation). In the cases treated in Sec. II, the starting points are rather the spectral decompositions of a signal and the frequency and amplitude modulations are derived. In Sec. II A, a signal that is a superposition of one strong and one weak signal of different frequencies is considered. The modulation frequency of the frequency and amplitude modulations is then given by the frequency difference between the two signals. The modulation index of the amplitude modulation and the frequency deviation of the frequency modulation are determined by the relative amplitudes of the weak and strong signals. The two modulations are of same phase for a weak-signal frequency above the strong-signal frequency, and of opposite phase for the reverse situation. The case of one strong signal and two weak signals is shown in Sec. II B to result in beating frequency and amplitude modulations. The modulation frequency is the average of the two frequency differences between the strong signal and the two weak sig-

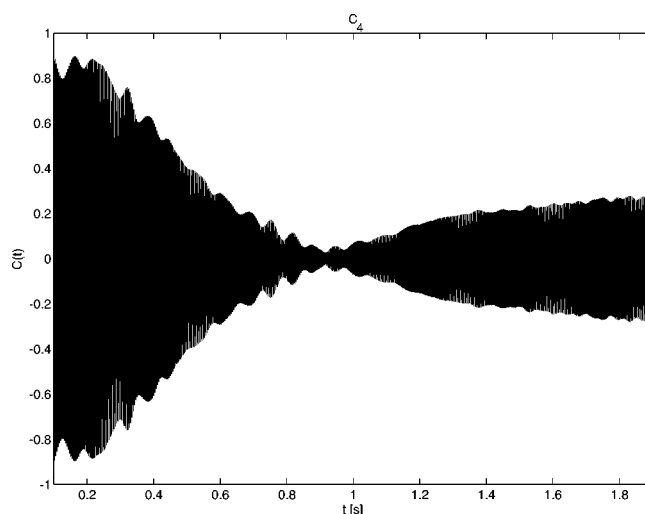


FIG. 13. Same as Fig. 6, but for the same sound as in Fig. 12.

nals. It is the difference between these two frequency differences that gives rise to the beating character of the frequency and amplitude modulations and that determines the beat frequency. For weak signals of equal amplitude, the modulation index of the amplitude modulation doubles and the frequency deviation of the frequency modulation approximately doubles, as compared with the case of one strong and one weak signal. The frequency and amplitude modulations have a relative phase shift near $\pm \pi/2$ in between the extrema of the beat envelopes. The relative phase shift between the beatings of the two modulations is π .

By digital recording of the sound of piano tones, beating frequency and amplitude modulations are observed in fair agreement with the derived influence of weakly excited adjacent tones one semitone step down and up, as described in Sec. III. Of immediate interest then is to what extent these phenomena are audible. Detection thresholds for both amplitude and frequency modulations have long been examined.^{32–38} For sinusoidal amplitude modulation, Zwicker^{33–35} studied the threshold value of the modulation index as a function of selected values of carrier frequency, modulation frequency, and sound-pressure level. For a carrier frequency of 1 kHz, a modulation frequency of 15 Hz and a sound-pressure level of 80 dB, the threshold modulation index is about 3%. For the beating amplitude modulation, treated in Sec. IIB, the maximal modulation index is given by $2a$, which is about 10% with the approximate value of a derived from observations described in Sec. IIIA and illustrated in Fig. 5. With respect to sinusoidal frequency modulation, Zwicker^{33–35} reported threshold values of the frequency deviation, again as a function of selected values of carrier frequency, modulation frequency, and sound-pressure level. For a carrier frequency of 250 Hz, a modulation frequency of 4 Hz and a sound-pressure level of 80 dB, the threshold frequency deviation is about 1.6 Hz, i.e., the threshold peak-to-peak frequency swing about 3.2 Hz. This is in fair agreement with a more recent study,³⁶ which also provides individual values for four subjects. For a carrier frequency of 250 Hz, a modulation frequency of 16 Hz and a sound-pressure level of 70 dB, the lowest detectable peak-to-peak frequency swing reported is 2.6 Hz. In frequency discrimination of pulsed pure tones, the detection thresholds are generally lower, about 1 Hz in the frequency range 200–400 Hz at a sound-pressure level of 40 dB.³⁷ From the observations described in Sec. IIIA and illustrated in Fig. 5, the maximal peak-to-peak frequency swing is about 3 Hz. It therefore appears that the audibility criterion is better fulfilled for the amplitude modulation than for the frequency modulation.

It has been observed that mixed amplitude and frequency modulation may be detected, although either of the modulations are below threshold.³⁸ The precise conditions of the two types of modulation are, however, important. In the case studied in the present work, the beating character of both frequency and amplitude modulations, with alternately occurring maxima of the beatings and an intricate phase relationship between the modulations, distinguishes the situation from those previously studied. The modulation and beat frequencies of frequency and amplitude modulation

also vary across the piano scale. The modulatory effects may also vary between different pianos due to different degrees of damping. Whereas the analysis in the present work refers to the fundamentals of tones, the sound of a tone derives from its full spectrum. The spectral composition of a piano tone is influenced by how hard strings are struck and the condition of the hammer.^{7–10,12,13,16,17} For all these reasons, further psychoacoustic investigations will be helpful in determining the possible contribution of the beating frequency and amplitude modulation deriving from coupling of tones to the quality of the piano tone.

The beating frequency and amplitude modulations both exceed the reported thresholds when the damping of the adjacent tones is released by means of the sustaining pedal, as can be concluded from Fig. 7 which refers to a situation with dampers lifted off the strings of the adjacent tones. The modulations can therefore be expected to contribute to the clearly audible difference in tone quality with and without the sustaining pedal being engaged, a difference which may also have other components. The dependence of the frequency and amplitude modulations on the detuning of strings in a unison may possibly serve an aural tuner as a complement to the principal tuning out the beats in obtaining an optimal tone quality. As the example in Fig. 10 demonstrates, even small changes of the detuning of unison strings may have large effects on the frequency modulation and cause large frequency deviations. This may also partly explain the observed preference for limited detunings of unison strings.²⁵ The frequency deviations observed in the present work are also large in comparison with the accuracy that an expert aural tuner can achieve. The latter accuracy has been estimated at ± 2 cents in the middle range of the piano scale,²⁴ which for C₄ is only about ± 0.3 Hz, to be compared with the observed maximal peak-to-peak frequency swing of about 3 Hz (Fig. 5). An even higher accuracy than ± 2 cents is actually required and achieved in tuning, e.g., equal-tempered fifths, for which the narrowing relative to beatless fifths in itself is approximately 2 cents. For electronically assisted tuning, high-resolution frequency meters are often based on the detection of successive zero crossings of the signal. An understanding of the phenomena discussed in the present work may therefore be of interest also in that context. The modulatory phenomena observed in a piano may occur in other stringed musical instruments transferring string vibrations to a soundboard via a bridge.

ACKNOWLEDGMENTS

Anders Askenfelt is gratefully acknowledged for valuable and stimulating discussions. Valuable advice from reviewers and the editor is also gratefully acknowledged.

¹H. von Helmholtz, *Die Lehre von Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, first edition 1862, English translation of the fourth edition 1885 by A. J. Ellis, *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, reprinted by Dover Publications, New York, 1954.

²*Five Lectures on the Acoustics of the Piano*, edited by A. Askenfelt (Royal Swedish Academy of Music, Stockholm, 1990). A www edition is available at www.speech.kth.se/music/5_lectures.

- ³N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*, 2nd ed. (Springer, Berlin, 1998).
- ⁴H. A. Conklin, Jr., "Design and tone in the mechanoacoustic piano. I. Piano hammers and tonal effects," *J. Acoust. Soc. Am.* **99**, 3286–3296 (1996).
- ⁵H. A. Conklin, Jr., "Design and tone in the mechanoacoustic piano. II. Piano structure," *J. Acoust. Soc. Am.* **100**, 695–708 (1996).
- ⁶H. A. Conklin, Jr., "Design and tone in the mechanoacoustic piano. III. Piano strings and scale design," *J. Acoust. Soc. Am.* **100**, 1286–1298 (1996).
- ⁷D. E. Hall, "Piano string excitation in the case of small hammer mass," *J. Acoust. Soc. Am.* **79**, 141–147 (1986).
- ⁸D. E. Hall, "Piano string excitation. II. General solution for a hard narrow hammer," *J. Acoust. Soc. Am.* **81**, 535–546 (1987).
- ⁹D. E. Hall, "Piano string excitation. III. General solution for a soft narrow hammer," *J. Acoust. Soc. Am.* **81**, 547–555 (1987).
- ¹⁰H. Suzuki, "Model analysis of hammer–string interaction," *J. Acoust. Soc. Am.* **82**, 1145–1151 (1987).
- ¹¹D. E. Hall and P. Clark, "Piano string excitation. IV. The question of the missing modes," *J. Acoust. Soc. Am.* **82**, 1913–1918 (1987).
- ¹²X. Boutillon, "Model for piano hammers: Experimental determination and digital simulation," *J. Acoust. Soc. Am.* **83**, 746–754 (1988).
- ¹³D. E. Hall and A. Askenfelt, "Piano string excitation. V. Spectra for real hammers and strings," *J. Acoust. Soc. Am.* **83**, 1627–1638 (1988).
- ¹⁴A. Askenfelt and E. Jansson, "From touch to string vibrations. I. Timing in the grand piano action," *J. Acoust. Soc. Am.* **88**, 52–63 (1990).
- ¹⁵A. Askenfelt and E. Jansson, "From touch to string vibrations. II. The motion of the key and hammer," *J. Acoust. Soc. Am.* **90**, 2383–2393 (1991).
- ¹⁶D. E. Hall, "Piano string excitation. VI. Nonlinear modeling," *J. Acoust. Soc. Am.* **92**, 95–105 (1992).
- ¹⁷A. Askenfelt and E. Jansson, "From touch to string vibrations. III. String motion and spectra," *J. Acoust. Soc. Am.* **93**, 2181–2196 (1993).
- ¹⁸O. L. Railsback, "A chromatic stroboscope," *J. Acoust. Soc. Am.* **9**, 37–42 (1937).
- ¹⁹O. L. Railsback, "Scale temperament as applied to piano tuning," *J. Acoust. Soc. Am.* **9**, 274 (1938).
- ²⁰O. L. Railsback, "A study of the tuning of pianos," *J. Acoust. Soc. Am.* **10**, 86 (1938).
- ²¹O. H. Schuck and R. W. Young, "Observations on the vibrations of piano strings," *J. Acoust. Soc. Am.* **15**, 1–11 (1943).
- ²²R. W. Young, "Inharmonicity of plain wire piano strings," *J. Acoust. Soc. Am.* **24**, 267–273 (1952).
- ²³D. W. Martin, "Decay rates of piano tones," *J. Acoust. Soc. Am.* **19**, 535–541 (1947).
- ²⁴D. W. Martin and W. D. Ward, "Subjective evaluation of musical scale temperament in pianos," *J. Acoust. Soc. Am.* **26**, 932 (1954); **33**, 582–585 (1961).
- ²⁵R. E. Kirk, "Tuning preferences for piano unison groups," *J. Acoust. Soc. Am.* **31**, 1644–1648 (1959).
- ²⁶G. Weinreich, "Coupled piano strings," *J. Acoust. Soc. Am.* **62**, 1474–1484 (1977).
- ²⁷G. Weinreich, "The coupled motions of piano strings," *Sci. Am.* **240**(1), 118–127 (1979).
- ²⁸G. Weinreich, "The coupled motion of piano strings," in Ref. 2.
- ²⁹B. Capleton, "False beats in coupled piano string unisons," *J. Acoust. Soc. Am.* **115**, 885–892 (2004).
- ³⁰W. M. Hartmann, *Signals, Sound, and Sensation* (Springer, New York, 1998).
- ³¹A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1989).
- ³²E. G. Shower and R. Biddulph, "Differential pitch sensitivity of the ear," *J. Acoust. Soc. Am.* **3**, 275–287 (1931).
- ³³E. Zwicker, "Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenzmodulation Eines Tones," *Acustica* **2**, AB 125–133 (1952).
- ³⁴E. Zwicker, "Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs," *Acustica* **6**, 365–381 (1956).
- ³⁵E. Zwicker and R. Feldtkeller, *Das Ohr als Nachrichtenempfänger*, 2nd ed. (Hirzel, Stuttgart, 1967).
- ³⁶L. Demany and C. Semal, "Detection thresholds for sinusoidal frequency modulation," *J. Acoust. Soc. Am.* **85**, 1295–1301 (1989).
- ³⁷C. C. Wier, W. Jesteadt, and D. M. Green, "Frequency discrimination as a function of frequency and sensation level," *J. Acoust. Soc. Am.* **61**, 178–184 (1977).
- ³⁸E. Ozimek and A. Sek, "Perception of amplitude and frequency modulated signals (mixed modulation)," *J. Acoust. Soc. Am.* **82**, 1598–1603 (1987).

Generation of longitudinal vibrations in piano strings: From physics to sound synthesis^{a)}

Balázs Bank^{b)} and László Sujbert

Department of Measurement and Information Systems, Budapest University of Technology and Economics, 1521 Budapest, Hungary

(Received 14 December 2004; revised 18 January 2005; accepted 19 January 2005)

Longitudinal vibration of piano strings greatly contributes to the distinctive character of low piano notes. In this paper a simplified modal model is developed, which describes the generation of phantom partials and longitudinal free modes jointly. The model is based on the simplification that the coupling from the transverse vibration to the longitudinal polarization is unidirectional. The modal formulation makes it possible to predict the prominent components of longitudinal vibration as a function of transverse modal frequencies. This provides a qualitative insight into the generation of longitudinal vibration, while the model is still capable of explaining the empirical results of earlier works. The semi-quantitative agreement with measurement results implies that the main source of phantom partials is the transverse to longitudinal coupling, while the string termination and the longitudinal to transverse coupling have only small influence. The results suggest that the longitudinal component of the tone can be treated as a quasi-harmonic spectrum with formantlike peaks at the longitudinal modal frequencies. The model is further simplified and applied for the real-time synthesis of piano sound with convincing sonic results. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1868212]

PACS numbers: 43.75.Mn, 43.75.Wx, 43.40.Cw [NHF]

Pages: 2268–2278

I. INTRODUCTION

In this paper the generation mechanism of longitudinal vibration in piano strings is investigated. The purpose of this paper is twofold: to explain the experimental results of earlier papers and to provide a guideline for physics-based sound synthesis.

The importance of longitudinal vibration of piano strings was recognized long ago by piano builders. Conklin¹ demonstrated that the pitch relation of the transverse and longitudinal component strongly influences the quality of the tone and described a method to tune these components. Giordano and Korty² found that the amplitude of the longitudinal vibration is a nonlinear function of the amplitude of transverse vibration, confirming the assumption that the longitudinal component is generated by the nonlinearity of the string and not by the “misalignment” of the hammer.

Nakamura and Naganuma³ found a second series of partials in piano sound spectra having one-fourth of inharmonicity compared to the main partial series. They attributed these to the horizontal polarization of the string, but they have actually found the series that later was named “phantom partials” by Conklin. Conklin⁴ pointed out that the phantom partials are generated by nonlinear mixing and their frequencies are the sum or difference of transverse model frequencies. He named “even phantoms” those having double the frequency ($2f_n$), of a transverse mode and “odd

phantoms” those which appear at the sum $f_m + f_n$ or difference $f_m - f_n$ frequencies of two transverse modes. Conklin’s measurements have shown that odd phantoms generally originate from adjacent parents, i.e., can be found at $f_5 + f_6$ rather than at $f_4 + f_7$. Phantom partials have also been found in the spectrum of guitar tones.⁴ In a recent paper about guitar transients, Woodhouse states that the amplitude of the phantom partials seems to be modulated according to the longitudinal modal frequencies.⁵ The present paper gives a theoretical explanation for these experimental results.

In an earlier work⁶ some of the properties of phantom partials and longitudinal modes have been investigated. It was pointed out and it is emphasized here again that longitudinal modes and phantom partials are two different manifestations of the same phenomenon: they are the free and forced response of the same system, respectively. Therefore, in the theoretical treatment of the present paper they are covered jointly. This paper outlines some of the findings of Ref. 6 and provides a more refined theoretical background.

Theoretical works on nonlinear string vibrations that consider longitudinal motion include the papers of Narasimha,⁷ Anand,⁸ Watzky,⁹ and O’Reilly and Holmes.¹⁰ As these papers discuss the nonlinear coupling of the first modes of the two transverse polarizations, the inertial effects of the longitudinal vibration are neglected. This leads to the uniformity of the tension and equals to computing the tension from the relative elongation of the string.⁸ (This is also studied in the Appendix.) The uniform tension approximation (thoroughly discussed by Legge and Fletcher¹¹) often forms the basis for the sound synthesis of nonlinear strings. Papers include tension-modulated string models^{12,13} based on digital waveguides¹⁴ and energy-conserving finite difference schemes.¹⁵ The inertial effects of longitudinal modes have

^{a)}Portions of this work have been presented in “Modeling the longitudinal vibration of piano strings,” Proceedings of the Stockholm Music Acoustics Conference, Stockholm, Sweden, August 2003, and “A piano model including longitudinal string vibrations,” Proceedings of the 7th International Digital Audio Effects Conference, Naples, Italy, October 2004.

^{b)}Electronic mail: bank@mit.bme.hu

been considered in the case of rubberlike strings by Leamy and Gottlieb¹⁶ and by Kurmyshev.¹⁷

However, the results of these earlier papers^{7-11,16,17} cannot be directly applied for the present purposes. This is because the present problem is more complex in the sense that not only the first few but 50 to 100 transverse modes have to be taken into account in the case of a struck piano string. On the other hand, it is simpler in the way that the longitudinal to transverse coupling and the coupling of different transverse modes are not investigated, as the primary interest is on the longitudinal vibration itself. Therefore, in this paper a modal model is developed that computes the spectrum of the longitudinal vibration in the case of arbitrary transverse model frequencies. It turns out that the most important properties of the longitudinal vibration can be explained by this simplified modal approach. The Appendix shows that the uniform tension approximation of earlier papers⁷⁻¹¹ is the special case of the tension computed from the modal model presented here.

The synthesis of the longitudinal components in piano tone is quite a recent topic. In an earlier work⁶ the digital waveguide string model¹⁴ has been extended by an auxiliary digital waveguide for computing the response of the phantom partials and by second-order resonators for modeling the longitudinal free modes. A very efficient method for modeling the phantom partials was proposed by Bensa.¹⁸ The model had a loose connection to physical reality, since in that work spatially uniform tension was assumed. In a recent work¹⁹ a physics-based solution is presented for modeling the longitudinal components jointly (i.e., phantom partials and free modes together), which is able to produce high quality piano sounds. In this model second-order resonators are nonlinearly excited according to the transverse string shape computed by a finite difference model. This paper outlines the above approach and presents an alternative technique having lower computational cost.

The paper is organized as follows: first the differential equations are derived from basic principles in Sec. II. In Sec. III a modal model is presented, which analytically computes the longitudinal vibration under certain assumptions. In Sec. IV the measurements of other authors²⁻⁴ are explained by the results of the proposed model, and some measurement results of the present authors are also given. Section V describes efficient algorithms for the synthesis of piano sound including the longitudinal vibrations. The Appendix relates the uniform tension approximation to the modal model presented in this paper.

II. EQUATIONS FOR ONE PLANE OF VIBRATION

A real piano string is vibrating in two transverse planes, and in the longitudinal direction as well. Principally, piano hammers excite one transverse polarization of the string, while the other two polarizations are gaining energy through coupling. For simplicity, it is assumed in this section that the string is vibrating in one plane, i.e., one transverse and one longitudinal polarization are present. Losses and dispersion are neglected for the time being and they are included later in Sec. III. The differences arising from the incorporation of two transverse polarizations are outlined in Sec. III E. Note

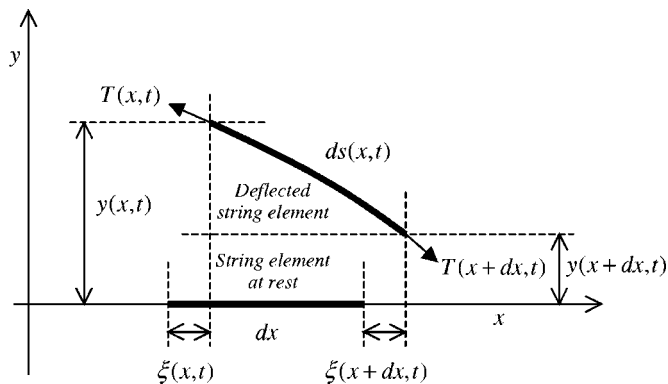


FIG. 1. The string element.

that similar derivations in a different formulation can be found in earlier works, e.g., in the textbook of Morse and Ingard.²⁰ The equations are developed here to make the paper self-explanatory.

When a transverse displacement occurs on the string, the string elongates. This results in a force exciting a longitudinal wave in the string. The longitudinal wave modulates the tension of the string, which influences the transverse vibration.

The element of length dx at equilibrium will have the length ds , as depicted in Fig. 1, which is calculated as follows:

$$ds^2(x,t) \approx [\xi(x+dx,t) - \xi(x,t) + dx]^2 + [y(x+dx,t) - y(x,t)]^2. \quad (1)$$

As dx is infinitesimally small, the differences are substituted by differentials

$$ds = \sqrt{\left(\frac{\partial \xi}{\partial x} + 1\right)^2 dx^2 + \left(\frac{\partial y}{\partial x}\right)^2 dx^2}, \quad (2)$$

where $y = y(x,t)$ and $\xi = \xi(x,t)$ are the transverse and longitudinal displacements of the string with respect to time t and space x . The tension $T = T(x,t)$ of the string (which equals to T_0 at rest) is calculated according to Hooke's law,

$$T = T_0 + ES \left(\frac{ds}{dx} - 1 \right), \quad (3)$$

where E is the Young's modulus and S is the cross-section area of the string. By substituting Eq. (2) into Eq. (3) and neglecting the higher order terms the tension can be approximated as

$$T \approx T_0 + ES \left[\frac{\partial \xi}{\partial x} + \frac{1}{2} \left(\frac{\partial y}{\partial x} \right)^2 \right]. \quad (4)$$

As the segment ds is nearly parallel to the x axis, the longitudinal force on the segment ds can be approximated as the difference of the tension at the sides of the segment

$$F_x \approx \frac{\partial T}{\partial x} dx \approx ES \left[\frac{\partial^2 \xi}{\partial x^2} + \frac{1}{2} \frac{\partial (\partial y / \partial x)^2}{\partial x} \right] dx. \quad (5)$$

Note that the resolution of the force would only introduce a negligible correction term for metal strings (where $T_0 \ll ES$). The force F_x acts on a mass μdx where μ is the mass

per unit length. Accordingly, the longitudinal vibration is approximately described by

$$\mu \frac{\partial^2 \xi}{\partial t^2} = ES \frac{\partial^2 \xi}{\partial x^2} + \frac{1}{2} ES \frac{\partial(\partial y/\partial x)^2}{\partial x}, \quad (6)$$

which is the standard one-dimensional wave equation with an additional force term depending on the transverse vibration. According to Eq. (6), the transverse string motion can only excite the longitudinal vibration if the square of the string slope is significant, i.e., the transverse displacement is relatively large. Note that neglecting the term $\mu(\partial^2 \xi/\partial t^2)$ leads to the uniform tension approximation (see the Appendix).

After similar derivations the wave equation for the transverse motion can be written as

$$\mu \frac{\partial^2 y}{\partial t^2} = T_0 \frac{\partial^2 y}{\partial x^2} + ES \frac{\partial\{(\partial y/\partial x)[\partial \xi/\partial x + \frac{1}{2}(\partial y/\partial x)^2]\}}{\partial x}, \quad (7)$$

which is again a one-dimensional wave equation with an additional force term depending on the product of the transverse slope and the tension variation. Consequently, the longitudinal vibration influences the transverse one if both the transverse and longitudinal displacements are relatively large.

To sum up, both Eqs. (6) and (7) can be considered as standard linear wave equations with additional forcing terms on their right-hand sides. From Eq. (6) it can be concluded that the level of transverse to longitudinal coupling depends on the magnitude of transverse vibration according to a square law. From Eq. (7) it turns out that the amount of longitudinal to transverse coupling is a third-order function of the amplitude of transverse vibration, since ξ is in the order of y^2 [see Eq. (6)].

III. TRANSVERSE TO LONGITUDINAL COUPLING

As the main interest of this paper is to clarify the generation of longitudinal string vibration, a further simplification is made: the longitudinal to transverse coupling is neglected. Moreover, the string termination is assumed to be infinitely rigid. The effects of these assumptions are covered in Sec. III F.

These limitations lead to a model that cannot be in complete quantitative agreement with measurements. On the other hand, its simplicity helps to gain a better understanding of the phenomenon. This simple model is already enough for the qualitative explanation of the measurements of earlier papers,²⁻⁴ as discussed in Sec. IV later. Furthermore, sound synthesis models (see Sec. V) based on these principles produce realistic piano sounds.

A. Excitation force

For the freely vibrating, dispersive, lossy, and rigidly terminated string the transverse displacement for a given position $0 \leq x \leq L$ and time $t \geq 0$ can be written in the following form:²¹

$$y(x, t) = \sum_{n=1}^{\infty} A_n \cos(2\pi f_n t + \varphi_n) e^{-t/\tau_n} \sin\left(\frac{n\pi x}{L}\right), \quad (8)$$

where f_n is the frequency, τ_n is the decay time, A_n is the initial amplitude, and φ_n is the initial phase of the transverse mode n , and L refers to the length of string. This form is of particular interest since the motion of the piano string behaves similarly after the hammer-string contact (typically after 1–2 ms).

As a general case, the transverse displacement of a rigidly terminated string can be described by a similar formula:

$$y(x, t) = \sum_{n=1}^{\infty} y_n(t) \sin\left(\frac{n\pi x}{L}\right), \quad (9)$$

where the time-dependent terms of Eq. (8) are substituted by the series of functions $y_n(t)$, which can be considered as the instantaneous amplitudes of the modes $n=1, \dots, \infty$. This notation is used to obtain simpler formulas in the following derivations.

According to Eq. (6), the transverse to longitudinal excitation-force distribution $F_{t \rightarrow l}(x, t)$ is computed as

$$F_{t \rightarrow l}(x, t) = \frac{1}{2} ES \frac{\partial[\partial y(x, t)/\partial x]^2}{\partial x}. \quad (10)$$

Substitution of Eq. (9) into Eq. (10) yields

$$F_{t \rightarrow l}(x, t) = \frac{1}{2} ES \frac{\partial[\sum_{n=1}^{\infty} y_n(t)(n\pi/L)\cos(n\pi x/L)]^2}{\partial x}, \quad (11)$$

which, after some derivations, takes the following form:

$$F_{t \rightarrow l}(x, t) = -ES \frac{\pi^3}{4L^3} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} y_m(t) y_n(t) mn \times \left[(m+n) \sin\left(\frac{m+n}{L} \pi x\right) + (m-n) \sin\left(\frac{m-n}{L} \pi x\right) \right]. \quad (12)$$

Note that the indices m and n belong to variables of transverse modes throughout the paper. The variables of longitudinal modes are indexed by k .

B. Longitudinal motion

The longitudinal displacement can be written in the same form as Eq. (9), resulting in

$$\xi(x, t) = \sum_{k=1}^{\infty} \xi_k(t) \sin\left(\frac{k\pi x}{L}\right). \quad (13)$$

By applying the derivations of Morse²² (which were originally developed for transverse vibrations), the instantaneous amplitude $\xi_k(t)$ of the longitudinal mode k is obtained as

$$\xi_k(t) = F_{t \rightarrow l, k}(t) * \xi_{\delta, k}(t), \quad (14a)$$

$$F_{t \rightarrow l, k}(t) = \int_0^L F_{t \rightarrow l}(x, t) \sin\left(\frac{k\pi x}{L}\right) dx, \quad (14b)$$

$$\xi_{\delta, k}(t) = \frac{1}{\pi L \mu} \frac{e^{-t/\tau_k'}}{f_k'} \sin(2\pi f_k' t), \quad (14c)$$

where the * sign denotes time-domain convolution and $F_{t \rightarrow 1,k}(t)$ is the excitation force acting on the longitudinal mode k . The time-domain impulse response of longitudinal mode k is denoted by $\xi_{\delta,k}(t)$, where f'_k and τ'_k stand for the frequency and decay time of the longitudinal mode k . Note that the single quote in f'_k and τ'_k is used to distinguish the longitudinal variables from the transverse ones.

The first step in calculating the longitudinal motion is the computation of the excitation force $F_{t \rightarrow l,k}(t)$ by Eq. (14b), which is the scalar product of the excitation-force distribution $F_{t \rightarrow l}(x,t)$ and the longitudinal modal shape. From Eqs. (12) and (14b) it follows that $F_{t \rightarrow l,k}(t)$ is nonzero for $m+n=k$ and $|m-n|=k$ only, since in all other cases the spatial distribution of the excitation $F_{t \rightarrow l}(x,t)$ is orthogonal to the modal shape of mode k .

The two cases can be computed separately by defining $F_{t \rightarrow l,k}(t)$ as a sum of two components, i.e., $F_{t \rightarrow l,k}(t) = F_{t \rightarrow l,k}(t)^+ + F_{t \rightarrow l,k}(t)^-$. The component originating from the transverse modes that satisfy $m+n=k$ is

$$F_{t \rightarrow 1,k}(t)^+ = -ES \frac{\pi^3}{8L^2} \sum_{n=1}^{k-1} y_{k-n}(t) y_n(t) k(k-n)n. \quad (15a)$$

The component coming from $|m-n|=k$ becomes

$$F_{t \rightarrow 1,k}(t)^- = -2ES \frac{\pi^3}{8L^2} \sum_{n=1}^{\infty} y_{k+n}(t) y_n(t) k(k+n)n. \quad (15b)$$

The factor of 2 in Eq. (15b) comes from the fact that there are two equal series $m=k+n$ and $n=k+m$, since both satisfy $|m-n|=k$.

C. Excitation frequencies

For qualitative understanding of the longitudinal components it is useful to look at the spectra of the excitation force series $F_{t \rightarrow l,k}(t)$. The most important question is where the frequency peaks can be found.

To the first approximation, the instantaneous amplitudes $y_n(t)$ are exponentially decaying sinusoidal functions with the frequencies f_n , such as in Eq. (8). By observing Eqs. (15a) and (15b) the frequencies of the mixing terms in $F_{t \rightarrow l,k}(t)$ can be calculated as

Frequencies

$$\begin{aligned} \text{in } F_{t \rightarrow l,k}(t)^+ : & \begin{cases} f_n + f_{k-n} \approx f_k, \\ f_n - f_{k-n} \approx f_{|2n-k|}, \end{cases} \\ \text{in } F_{t \rightarrow l,k}(t)^- : & \begin{cases} f_n + f_{k+n} \approx f_{2n+k}, \\ f_n - f_{k+n} \approx f_k, \end{cases} \end{aligned} \quad (16)$$

where the form f_a refers to the frequency of the transverse mode with mode number a .

1. Harmonic transverse vibration

The approximations in Eq. (16) become equalities if the transverse frequencies f_n are perfectly harmonic, i.e., $f_n = n f_0$, which is the case for string instruments having negligible inharmonicity. In this case, there is a strong peak at the frequency f_k , and a series of partials at f_{2n-1} for odd k and at f_{2n-2} for even k , with $n = 1, \dots, \infty$. This means that the

odd longitudinal modes are excited by components having the same frequencies as the odd transverse modes. Similarly, the even longitudinal modes are excited at the even transverse modal frequencies.

These frequencies form the inputs of the time-domain impulse responses $\xi_{\delta,k}(t)$, which can be considered as second-order resonators [see Eq. (14c)]. The output of a resonator has two types of components: one component is the free response, which is a decaying sinusoid at the frequency f'_k . The other component is the forced response consisting of the frequency series f_{2n-1} or f_{2n-2} with $n = 1, \dots, \infty$. The amplitudes of these spectral lines are amplified around the peak of the resonator f'_k . As the responses of all the longitudinal modes are summed together, the output becomes similar to having formants on a rich harmonic spectrum. In any case, the forced components are indistinguishable from the transverse ones since they are exactly at the same frequencies.

2. Inharmonic transverse vibration

In the case of the piano, due to the stiffness of the string, the transverse partial frequencies do not form a perfect harmonic series but rather obey the equation²³

$$f_n = f_0 n \sqrt{1 + B n^2}, \quad (17)$$

where B is the inharmonicity coefficient and $f_0 \approx f_1$ is the fundamental frequency of the string.

In this case the terms $f_n + f_{k-n}$ and $f_n - f_{k+n}$ do not have the frequency of f_k but form a bunch of peaks around f_k . The peaks at the frequencies $f_n - f_{k-n}$ lie somewhat higher compared to $f_{|2n-k|}$ and the frequencies $f_n + f_{k+n}$ are lower than f_{2n+k} . This means that these peaks depart from the transverse partials in a rate determined by the inharmonicity coefficient B and the longitudinal mode number k . However, it is still true that odd longitudinal modes are excited by an oddlike partial series, while even longitudinal modes are excited by an evenlike one.

The force exciting the first longitudinal mode $F_{t \rightarrow 1,1}(t)$ is displayed in Fig. 2(a) by a solid line, computed by the resonator-based string model (see Sec. VC), which is the discrete-time implementation of the modal model described in Secs. III A and III B. Note that the excitation force has an oddlike partial series. The spectrum of the transverse bridge force is displayed by dots to show the transverse modal frequencies as a reference. The dashed line indicates the Fourier transform of the impulse response of the first longitudinal mode $\xi_{\delta,1}(t)$, amplifying the frequencies around 690 Hz. Figure 2(b) shows the excitation-force spectrum of the second longitudinal mode for the same example. It can be seen that here the excitation spectrum contains even partials only and that the peak of the longitudinal mode (dashed line) is located at a higher frequency (1380 Hz in this case).

The longitudinal motion is the sum of the motion of different modes. This means that spectra similar to Figs. 2(a) and (b) should be superimposed with slightly shifted excitation frequencies and very different longitudinal modal frequencies. The result is similar to formants on a quasi-harmonic spectrum but here the peaks are somewhat smeared as they are made up of many close frequencies. The most

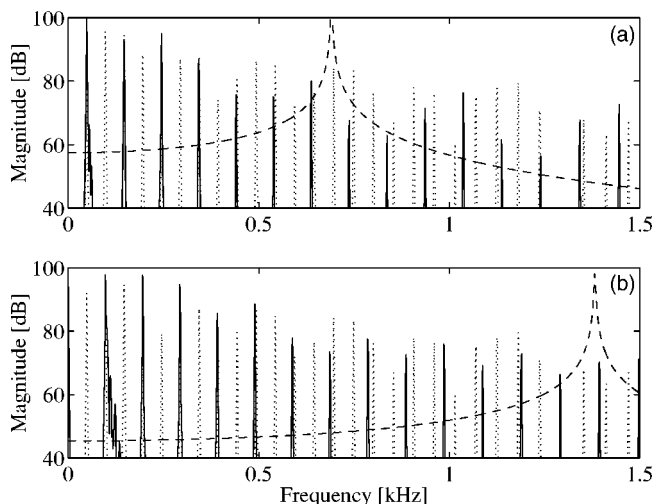


FIG. 2. The force spectrum exciting the first (a) and the second (b) longitudinal modes $[F_{t \rightarrow l,1}(t)$ and $F_{t \rightarrow l,2}(t)]$ computed by the resonator-based string model of Sec. V C (displayed by solid line). The transverse bridge force (dotted line) is displayed to show the transverse modal frequencies. The dashed line shows the frequency response of the first (a) and the second (b) longitudinal modes. The relative levels of the signals are arbitrary.

important difference from the case of harmonic transverse vibration is that these smeared peaks appear between the transverse ones and therefore they can be easily distinguished.

D. Longitudinal force at the bridge

The longitudinal component of the tone is transmitted to the soundboard via the force acting on the bridge in the longitudinal direction. This force equals to the tension [see Eq. (4)] at the termination $x=L$, written as

$$F_l(t) = - \left[T_0 + ES \frac{\partial \xi}{\partial x} \Big|_{x=L} + \frac{1}{2} ES \left(\frac{\partial y}{\partial x} \Big|_{x=L} \right)^2 \right], \quad (18)$$

showing that the force $F_l(t)$ depends not only on the longitudinal motion but on the transverse vibration as well.

The force component coming from the transverse motion have the same sum- and difference-frequency terms as the component arising from the longitudinal motion, but their amplitudes are different. It is an interesting outcome that when the transverse motion of the string contains low frequency components only, most of these terms cancel out and only the double frequency terms remain. This produces the same longitudinal bridge force as what would occur with applying the uniform tension approximation.¹¹ The derivation of this result is included in the Appendix.

E. Extension to two transverse planes

Real strings vibrate in two transverse polarizations. The modal frequencies for these polarizations can be different for the same modes, mostly because of the direction-dependent termination impedance. This produces beating and two-stage decay in piano sound.²⁴ Working out the equations (1)–(6) for the three-dimensional case gives

$$\mu \frac{\partial^2 \xi}{\partial t^2} = ES \frac{\partial^2 \xi}{\partial x^2} + \frac{1}{2} ES \left[\frac{\partial(\partial y / \partial x)^2}{\partial x} + \frac{\partial(\partial z / \partial x)^2}{\partial x} \right], \quad (19)$$

where z is the string displacement in the direction perpendicular to the already considered transverse y and longitudinal x directions.

It follows from Eq. (19) that the excitation-force distribution $F_{t \rightarrow l}(x, t)$ is the superposition of the excitation-force distributions computed for the two transverse planes separately. Accordingly, if two modes vibrate in two planes perpendicular to each other, their sum and difference frequencies do not appear in the excitation force. In reality the vibrating planes are not perfectly perpendicular to each other (the motion is actually not even planar^{7–10}), meaning that mode m vibrating in one plane will mix with mode n vibrating in a different plane.

The modal frequencies $f_{n,1}$ and $f_{n,2}$ of the two transverse polarizations are slightly different for the same mode number n . Accordingly, the excitation components coming from the two transverse modes m and n consist of four different frequencies. For example, the sum-frequency components have the frequencies $f_{m,1} + f_{n,1}$, $f_{m,1} + f_{n,2}$, $f_{m,2} + f_{n,1}$, and $f_{m,2} + f_{n,2}$. The difference-frequency components can be expressed similarly.

F. Validity of the approximation

1. Neglecting the longitudinal to transverse coupling

The assumption of neglecting the longitudinal to transverse coupling is valid until the longitudinal vibration is small compared to the transverse one. However, if one of the excitation frequencies of the longitudinal mode k [see Eq. (16)] is very close to the resonant frequency f'_k of that mode, the longitudinal motion can have extremely large amplitude. This would not happen in reality since the longitudinal motion would diminish the amplitude of those transverse modes from which it originates (the total energy of transverse and longitudinal vibrations cannot increase). This stabilizing effect is not included in our modal model. On the other hand, these coincidences have a small practical significance from the sound synthesis viewpoint since they produce an unpleasant ringing sound even when computed by a finite difference string model (see Sec. V A) having bidirectional coupling. (This fact also implies that these annoying coincidences should also be avoided in real pianos by careful string- and scale design.)

The longitudinal to transverse coupling would also introduce some terms of third order in the amplitude of the transverse vibration, but their contribution is less significant compared to the second-order terms discussed here. To sum up, the frequencies predicted by the model of Sec. III should be in quantitative agreement with the dominant peaks found in real piano spectrum. The amplitude behavior is described properly for those peaks that do not coincide with the resonant frequency of the excited longitudinal mode. (This holds for most of the peaks.)

2. Perfectly rigid termination

The termination of piano strings is not perfectly rigid, contrary to the assumptions made here. As the impedance of the bridge is ca. 1000 times larger compared to the impedance of the string, its main effect is a change in the partial

frequencies f_n and decay time τ_n , which can be easily incorporated in Eq. (8). The modal shapes also change slightly [L is substituted by $L + \delta L$ in Eq. (8), see Ref. 11], meaning that none of the longitudinal modal shapes are completely orthogonal to the modal shapes of the excitation-force distribution. However, it is still true that the dominant force components are those computed by Eqs. (15a) and (15b). This is confirmed by finite difference simulations showing only a small change in the output when a more realistic termination model is applied.

The termination of piano strings can also contribute to the energy transfer between the transverse and longitudinal motion.¹¹ As such a coupling is linear, it does not introduce new terms by itself. The transverse frequencies can appear in the longitudinal motion, and, conversely, the longitudinal frequencies may turn up in the transverse vibration. However, those transverse and longitudinal components that have the same frequency cannot be distinguished in the sound pressure. The coupling through the bridge in combination with the transverse to longitudinal coupling along the string could produce new terms, but they are of fourth order in the amplitude of the transverse vibration.

IV. CONNECTIONS TO MEASUREMENTS

In this section the results of Sec. III are related to the measurements of other authors.²⁻⁴ On the one hand, this confirms the modal model presented in Sec. III. On the other hand, it helps to understand the theoretical reasons underlying the findings of these experimental studies.²⁻⁴

A. Parentage of phantom partials

From the theoretical point of view, phantom partials are the forced motion of longitudinal vibrations. An interesting property of odd phantom partials discovered by Conklin⁴ is that they originate from adjacent parents, i.e., they can be found at frequencies $f_m + f_n$ where $m - n = 1$.

By looking at $F_{t \rightarrow l, k}(t)$ in Eq. (16) it turns out that the frequencies $f_n + f_m = f_n + f_{k+n}$ are quite close to each other for $m + n = p$ (they would actually coincide in the case of a perfectly harmonic transverse vibration having the frequency f_{2n+k}). The question is which $f_m + f_n$ combination has the largest amplitude in the resulting sound. It follows from Eq. (16) that the different $f_m + f_n$ components belonging to the same smeared peak (i.e., $m + n = p$) excite different longitudinal modes. Namely, the frequency $f_m + f_n$ excites the longitudinal mode having the mode number $k = m - n$. Accordingly, that $f_m + f_n$ component results in the largest longitudinal motion which excites the longitudinal mode having a modal frequency f'_k close to the frequency $f_m + f_n$. In other words, if the frequency of a phantom partial group is close to the frequency f'_k of the k th longitudinal mode, it mainly originates from parents having mode number difference of k .

The lower odd phantom partials, which were measured by Conklin,⁴ most probably have frequencies to which the first longitudinal mode is the nearest. In this case the $f_m + f_n$ terms satisfying $m - n = 1$ dominate, which actually originate from adjacent parents f_n and f_{n+1} .

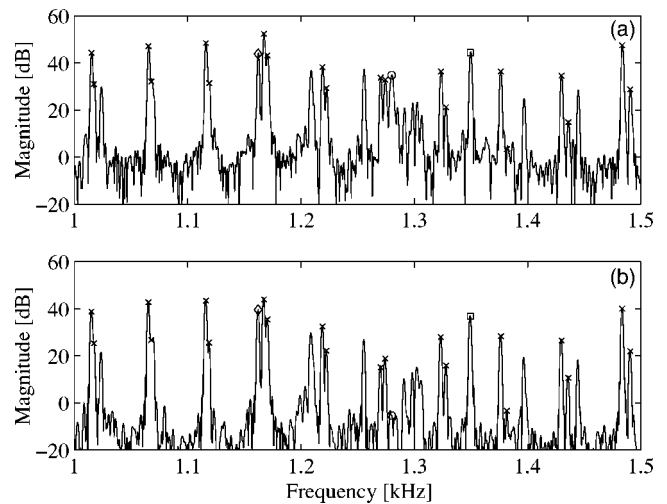


FIG. 3. Spectrum of the first (a) and the second second (b) of a F_1 piano tone. Transverse partials are marked by crosses and the second longitudinal mode is marked by a circle. Two prominent phantom partial groups are indicated by a square and a diamond (the latter is magnified in Fig. 4).

Similar considerations apply for even phantoms, that is, they are generated by parents having mode number difference of 2, 4, 6, etc., depending on the frequency of the phantom partial. However, there is an important difference that double frequency terms $2f_n$ also occur in the spectrum. These $2f_n$ components would arise even when the bandwidth of transverse components was significantly lower than the frequency of the first longitudinal mode, i.e., when the tension was approximately uniform along the string (see the Appendix). In other words, these are the only components that can be explained by the uniform tension approximation of Refs. 7–11, while for the sum- and difference frequency components the inertial effects of longitudinal modes have to be included in the model.

The spectrum of a recorded F_1 piano tone (having only one string) is displayed in Fig. 3. Transverse partials are indicated by crosses and the free response of the second longitudinal mode is marked by a circle. The remaining peaks are the forced response of the longitudinal motion, i.e., phantom partials. Figure 3(a) shows the first second of the tone and Fig. 3(b) displays the second, giving an insight to the evolution of the spectrum. Note that the free response of the longitudinal mode (circle) disappears fast in the noise (the decay time is ca. 0.15 s), while the phantom partials remain significant and their decay rate is comparable to that of the transverse partials. It can be said in general that the highest nontransverse peaks in the long-term spectrum are phantom partials amplified by a longitudinal mode (one prominent example is marked by a square). This suggests that the forced response of the longitudinal motion may have a larger perceptual significance than the free response itself. Most probably the pitch of the longitudinal component is determined by these amplified phantom partials (like the one marked by a square in Fig. 3) and not by the fast decaying free response. The interested reader may listen to the sound examples demonstrating the relative significance of these components.²⁵

The “single” phantom partial marked by a diamond in

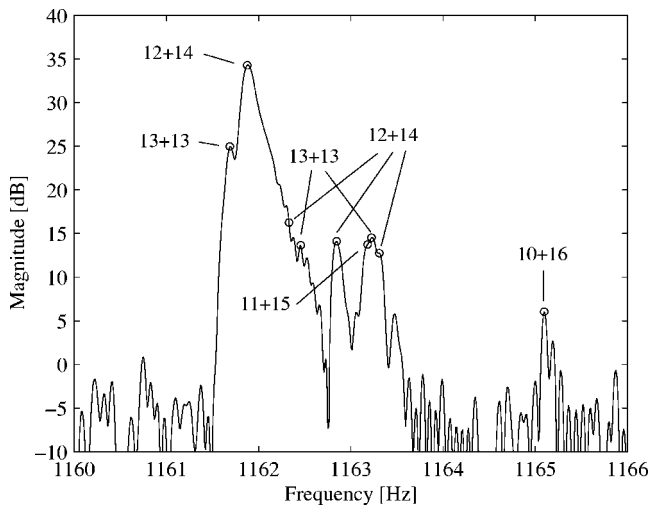


FIG. 4. The spectrum of an even phantom partial group in the F_1 piano tone of Fig. 3. Sum frequencies of transverse modes $f_m + f_n$ are marked by circles, and the mode numbers of the parent partials are labeled in the form of $m + n$. The phantom group is displayed by a diamond in Fig. 3.

Fig. 3 becomes a group of partials when plotted at a higher frequency resolution in Fig. 4. In this case the data length is 16 s (705 600 samples at $f_s = 44.1$ kHz), which was zero padded to 2^{22} samples after applying a Hanning window. The most prominent peaks of the phantom group are marked by circles. The label “ $m + n$ ” beside a circle indicate that the circle is located at the sum frequency of the transverse modes m and n (i.e., at $f_m + f_n$). The frequencies of the transverse modes were determined by finding peaks in the spectrum. Note that the same $m + n$ combinations can be found at several peaks: the reason is that the two different frequencies $f_{m,1}$ and $f_{m,2}$ of the two transverse polarizations of mode m mix with the two different frequencies $f_{n,1}$ and $f_{n,2}$ of mode n , as predicted in Sec. III E.

It can be seen in Fig. 4 that the highest peak comes from the 12th and 14th transverse modes and not from the 13th mode itself, although the amplitude of the latter is only 10 dB smaller. Other even phantoms show the same phenomenon: they principally originate from parents having mode-number difference of 2, 4, etc., and not from a single mode by frequency doubling. This contradicts the findings of Conklin⁴ but confirms the analysis of Sec. III.

B. Inharmonicity of phantom partials

Nakamura and Naganuma³ found that the inharmonicity of phantom partials (called “lower series” in Ref. 3) is one-fourth of that of normal transverse partials.

This can be explained by knowing that phantom partials are mainly produced by parents with mode numbers close to each other. This means that even phantoms have an approximate frequency of $f_p = 2f_n$, where $p = 2n$ is the “mode number” of the phantom partial. (See Fig. 4 as an example, where $f_{12} + f_{14} \approx 2f_{13}$.) Writing $f_p = 2f_n$ according to Eq. (17) and expressing the frequencies by the phantom mode number $p = 2n$ gives⁶

$$f_p \approx 2f_n = 2f_0n\sqrt{1 + Bn^2} = f_0p\sqrt{1 + \frac{1}{4}Bp^2}. \quad (20)$$

For even phantoms, the expression is quite accurate. For odd phantoms, $n = p/2$ is not an integer number. However, as the inharmonicity curve is a smooth function, the frequencies of odd phantom partials are also close to the ones predicted by Eq. (20).⁶

C. Amplitude of longitudinal vibration

Giordano and Korty² found that the amplitude of the longitudinal vibration is a nonlinear function of the amplitude of the transverse one. They noted that the nonlinear curve is faster than a simple quadratic function.

Equation (12) shows that a peak in the excitation spectrum of a longitudinal mode is a quadratic function of the overall amplitude of the generating transverse modes m and n . However, the amplitude of longitudinal motion is mostly determined by parents having sum frequencies $f_m + f_n$ around the longitudinal modal frequencies f_k' . The amplitude of these parents (with mode numbers around 10–20 in practice) are a nonlinear function of the overall amplitude of the transverse vibration. This is because of the nonlinear nature of the hammer–string interaction (see, e.g., Ref. 26). The presence of this second kind of nonlinearity explains why Giordano and Korty² could not measure a second-order relationship.

V. SOUND SYNTHESIS

The original motivation of this research was to support the development of physics-based piano models. These physics-based models do not have to describe each part of the instrument precisely. They should be as simple as possible while still producing agreeable sound quality. Therefore these models are often constructed from precise physical descriptions by neglecting those effects that have small perceptual relevance.

A. Finite difference modeling

A straightforward approach of modeling the vibration of piano strings is implementing the simultaneous differential equations (6) and (7) by the finite difference approach. Naturally, these have to be extended by the terms realizing frequency-dependent losses and dispersion. In an earlier work⁶ such a model was developed along the lines of the transverse string model of Chaigne and Askenfelt.^{27,28} However, the computational demand of such an approach is large because high sampling frequency ($f_s \approx 500$ kHz) is required due to the higher propagation speed in the longitudinal direction. Still, this approach can be very useful for experimental purposes. A commercial computer program based on a finite difference string model was written by Bernhard,²⁹ helping piano tuners in scale design.

The inclusion of longitudinal components in the piano model greatly improves the quality of synthesized piano sounds. However, complete finite difference modeling of the string is too demanding for real-time sound synthesis applications. In the following sections two models are presented that overcome this limitation. The composite model of Sec. VB replaces the finite difference model implementing Eq. (6) with the modal description for the longitudinal polariza-

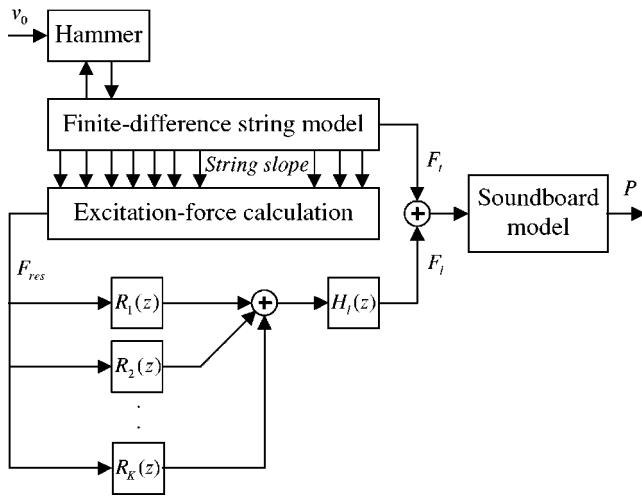


FIG. 5. The composite string model applying finite difference modeling and second-order resonators R_1, \dots, R_K .

tion, while the transverse vibration is computed by a variation of Eq. (7). The resonator-based string model of Sec. V C applies the modal model for both polarizations instead of directly realizing Eqs. (6) and (7).

B. The composite string model

In an earlier work¹⁹ a composite string model was introduced, which computes the longitudinal vibration in a modal form according to the theory presented in Sec. III. The model structure is depicted in Fig. 5. The transverse deflection $y(x, t)$ is computed by a finite difference string model running at audio sampling rate (e.g., $f_s = 44.1$ kHz), which implements the following differential equation:

$$\mu \frac{\partial^2 y}{\partial t^2} = T_0 \frac{\partial^2 y}{\partial x^2} - ES\kappa^2 \frac{\partial^4 y}{\partial x^4} - 2b_1\mu \frac{\partial y}{\partial t} + 2b_2\mu \frac{\partial^3 y}{\partial x^2 \partial t}. \quad (21)$$

Equation (21) is a variation of Eq. (7) where the nonlinear forcing term [the rightmost term of Eq. (7)] is missing, as the longitudinal to transverse coupling is neglected, while it is extended by terms realizing dispersion and losses. This is basically the string model proposed by Chaigne and Askenfelt,^{27,28} except for the last term, which substitutes temporal derivatives with spatial ones. This modification is suggested by Bensa *et al.*,³⁰ leading to a stable system for arbitrary b_2 . The κ parameter in the dispersion term refers to the radius of gyration of the string, and the constants b_1 and b_2 determine the decay rates of the partials. A finite difference hammer model^{27,28} is also attached to the string. The initial velocity of the hammer is denoted by v_0 in Fig. 5.

The excitation-force distribution of the longitudinal motion $F_{\rightarrow l}(x, t)$ is computed according to Eq. (10) from this transverse displacement. Then the excitation force $F_{\rightarrow l, k}(t)$ of each longitudinal mode k is computed by a scalar product with the longitudinal modal shape [see Eq. (14b)]. The instantaneous amplitudes $\xi_k(t)$ of the longitudinal modes are calculated according to Eq. (14a), which is implemented by second-order resonators (R_1, \dots, R_K in Fig. 5).

In order to reduce the computational cost, the same excitation force is used for all the longitudinal modes.¹⁹ This is

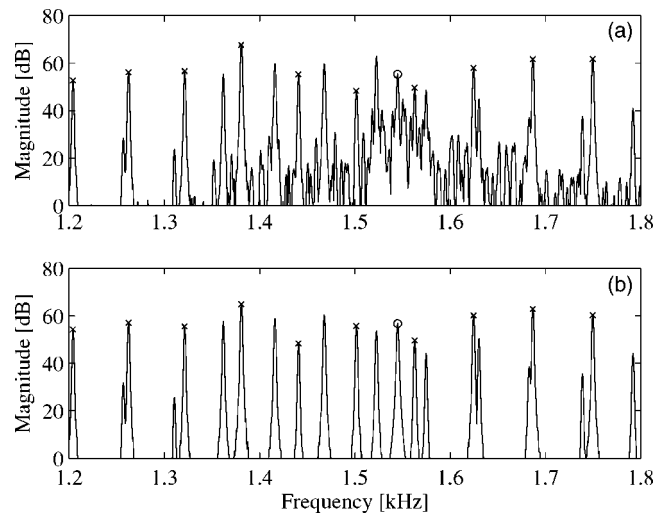


FIG. 6. The sound pressure spectrum of the first second of a synthesized G_1 piano tone²⁵ computed by the composite string model of Sec. V B (a) and by the resonator-based string model of Sec. V C (b). Crosses indicate the transverse partials and the second longitudinal mode is marked by a circle in both figures. To be compared with Fig. 3(a).

acceptable from a perceptual point of view because the spectra of the excitation forces $F_{\rightarrow l, k}(t)$ are quite similar. It follows from Eq. (16) that odd phantoms arise from the vibration of odd longitudinal modes and even phantoms from the vibration of even ones. Therefore it is satisfactory to compute the excitation force of one odd and one even longitudinal mode, e.g., the input of the resonators R_1, \dots, R_K can be $F_{\text{res}} = F_{\rightarrow l, 5}(t) + F_{\rightarrow l, 6}(t)$. This way a multiple (or smeared) peak of a phantom partial is substituted by a single, exponentially decaying sinusoid. In order to avoid an unpleasant ringing sound (see Sec. III F 1), the frequencies f'_k of the resonators R_1, \dots, R_K are set in a way that they do not coincide with the peaks of their excitation signal F_{res} . It has been found that using a common excitation force for R_1, \dots, R_K does not impair sound quality, but contributes to large computational savings.

The force signals F_t and F_l in Fig. 5 coming from the transverse and longitudinal polarizations are sent to the soundboard model, which computes the sound pressure P . The soundboard is modeled by a multi-rate filtering algorithm approximating the measured impulse response of a transversely excited piano soundboard.¹⁹ The soundboard responds differently to a longitudinal force than to a transverse one. This difference is modeled by a simple high-pass filter $H_l(z)$ in the longitudinal force path.

The sound pressure spectrum of the first second of a synthesized G_1 note²⁵ is displayed in Fig. 6(a). The phantom partials are clearly visible between the transverse modes, which are emphasized around the longitudinal free mode at 1450 Hz. The circle indicates the component coming from the longitudinal free response, while the crosses show the transverse modal frequencies. It can be seen that the spectrum is similar to that of a real piano tone displayed in Fig. 3. The composite string model produces the same sound quality as the full finite difference method of Sec. V A, while its computational requirements are reduced by an order of a magnitude (to around 10%–15%).

C. The resonator-based string model

The resonator-based string model is the discretization of the equations presented in Sec. III. The string displacement is represented by its modal form [see Eqs. (9) and (13)] for both the transverse and longitudinal polarizations and the instantaneous amplitudes $y_n(t)$ and $\xi_k(t)$ are computed by second-order resonators.

The string is excited by a hammer in the transverse polarization. The hammer is modeled in the same way as in the case of finite difference string models.^{27,28} The string response to the hammer force is calculated by a set of second-order resonators, which have input and output coefficients depending on the hammer position. The outputs of these resonators correspond to the instantaneous amplitudes $y_n(t)$ of the transverse vibration, which can be directly used to compute the excitation force $F_{t \rightarrow l, k}(t) = F_{t \rightarrow l, k}(t)^+ + F_{t \rightarrow l, k}(t)^-$ of the longitudinal modes by using Eqs. (15a) and (15b). From this point, the approach is the same as taken in Sec. VB: the excitation force of one even and one odd longitudinal mode is calculated and summed [e.g., $F_{\text{res}} = F_{t \rightarrow l, 5}(t) + F_{t \rightarrow l, 6}(t)$]. This signal is then fed to the resonators calculating the instantaneous amplitudes $\xi_k(t)$ of the longitudinal modes. The efficiency can be further increased if those components of the excitation signal F_{res} are not computed where the gain of the longitudinal resonator bank is small.

This model is capable of producing the same sound quality as the model of Sec. VB when the number of resonators implementing the transverse modes equals to the number of string elements in the finite difference model. Figure 6(b) displays the sound pressure spectrum of the first second of a G_1 piano tone synthesized by the resonator-based string model.²⁵ It can be seen in Fig. 6 that the resonator-based model produces a similar output compared to the composite model of Sec. VB when the string and hammer parameters are set to be the same. The only difference is that the composite string model generates noiselike peaks between the dominant partials due to computational inaccuracies. However, this is not considered as an advantage because the difference between the output of the two models is almost inaudible.²⁵

An advantage of this approach is that the computational complexity is reduced to less than the half. Moreover, this method is particularly advantageous when the goal is to reproduce a tone which is similar to that of a particular piano since the measured partial frequencies f_n and decay times τ_n can be directly implemented in the model. On the other hand, the resonator-based model is less physical in the sense that the physical parameters of the string (such as string mass and tension) have only indirect connection to the model.

D. Implications to psychoacoustic research

The informal listening tests made during the development of sound synthesis algorithms raised some questions on the perceptual aspects of longitudinal vibrations. It is an important property of the piano sound that the longitudinal component sounds as an inherent part of the tone, while it is still possible to perceive its pitch.²⁵ The present authors be-

lieve that the reason for this is that the distance of the phantom partials reinforces the pitch information originating from the transverse vibration, while the distance of the “formant peaks” leads to the pitch perception of the longitudinal mode. The perceptual effect can be similar to the sound of Touvinian throat- or overtone singers³¹ who can produce two tones simultaneously. Listening tests should be conducted to confirm this conjecture.

VI. CONCLUSIONS

The generation of longitudinal components can be summarized as follows: the longitudinal motion is continuously excited by the transverse vibration along the string (and not only during the hammer-string contact). The forced response to this excitation gives a rise to phantom partials, while the free response produces the components corresponding to the longitudinal modal frequencies. Both of these components develop under the assumption of rigid string terminations, i.e., the piano bridge has a less significant effect on the phenomenon.

According to the modal model presented in Sec. III, each longitudinal mode can be viewed as a second-order resonator, whose input is a quasi-harmonic spectrum, containing terms with sum and difference frequencies of some specific transverse modes. As each longitudinal mode emphasizes the peaks around its modal frequency, the sum of their outputs is similar to having formants on a quasi-harmonic spectrum.

In Sec. IV the experimental results of earlier papers have been explained by the results of the modal model, such as why phantom partials originate from adjacent parents and what the inharmonicity coefficient of the phantom partial series is. Some measurements of the present authors have also been outlined, confirming the results of the theoretical model.

Based on the further simplification of the model, two sound synthesis algorithms have been presented in Sec. V. The first one computes the transverse vibration by a finite difference string model and then calculates the inputs of the resonators, which represent the longitudinal modes. The second approach computes both the transverse and longitudinal vibrations in the modal domain, implemented by second-order resonators. Both models produce convincing piano sounds.²⁵

The present work has examined the main effects arising from the longitudinal vibration of piano strings. Secondary effects coming from the coupling of longitudinal to transverse polarizations and originating from the characteristics of the string termination could be a subject of future research. For these studies the admittance matrix measurement of the piano bridge would be of great significance. Moreover, the radiation properties of the soundboard as a function of longitudinal string force could also be an interesting field of research. As for sound synthesis models, the computational complexity might be lowered by further simplifications based on psychoacoustic criteria. Accordingly, studies on how the longitudinal components are perceived would also be of great importance.

The findings of the paper can be useful not only in the field of piano acoustics but for researchers interested in the analysis or synthesis of other stringed instruments, such as the guitar. Understanding the generation mechanism of longitudinal components (e.g., which transverse partials excite a specific longitudinal mode) can help piano or guitar builders to achieve a better control over the nature of the tone. For the sound synthesis of other string instruments, the tension computed by the modal model might be used to improve the performance of synthesis models that presently apply the uniform tension approximation.

ACKNOWLEDGMENTS

The authors are grateful for the helpful comments of C. Erkut, F. Fontana, J. Pakarinen, and V. Välimäki.

APPENDIX: UNIFORM TENSION AS A SPECIAL CASE

The assumption of spatially uniform tension is often applied in the literature (see the work of Legge and Fletcher¹¹ or Refs. 7–10), since in that case the tension can be computed from the relative elongation of the string by

$$\bar{T}(t) = T_0 + ES \frac{1}{2L} \int_{x=0}^L \left(\frac{\partial y}{\partial x} \right)^2 dx, \quad (\text{A1})$$

where the string tension $T(x, t) = \bar{T}(t)$ is independent of position x . This assumption is based on the fact that when the speed of the longitudinal waves is much larger than that of the transverse ones, the longitudinal inertial effects can be neglected.⁸

In this case the tension contains terms which have double the frequencies of the corresponding transverse modes.¹¹ If the transverse displacement is written in a modal form as in Eq. (9), the tension $\bar{T}(t)$ is obtained as

$$\bar{T}(t) = T_0 + ES \frac{\pi^2}{4L^2} \sum_{n=1}^{\infty} y_n^2(t) n^2. \quad (\text{A2})$$

It is of some interest to see how the results of this formula can be developed as a special case of the modal model presented in this paper. This both confirms the modal model described in Sec. III and helps in finding the limits where the assumption of uniform tension can be applied.

It can be seen in Eq. (4) that the expression of the tension $T(x, t)$ is made up of three terms. By defining T_0 as the tension at rest, $T_l(x, t)$ as the tension component proportional to the longitudinal slope, and $T_t(x, t)$ as the tension component proportional to the square of the transverse slope, the total tension can be written as

$$T(x, t) = T_0 + T_l(x, t) + T_t(x, t). \quad (\text{A3})$$

If the transverse displacement is expressed in the modal form of Eq. (9), the tension component coming from the transverse slope becomes

$$\begin{aligned} T_t(x, t) &= ES \frac{1}{2} \left[\frac{\pi}{L} \sum_{n=1}^{\infty} y_n(t) n \cos\left(\frac{n\pi x}{L}\right) \right]^2 \\ &= ES \frac{\pi^2}{4L^2} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} y_m(t) y_n(t) mn \\ &\quad \times \left[\cos\left(\frac{m+n}{L} \pi x\right) + \cos\left(\frac{m-n}{L} \pi x\right) \right]. \quad (\text{A4}) \end{aligned}$$

The Laplace transform of the time domain impulse response $\xi_{\delta, k}(t)$ in Eq. (14c) of a longitudinal mode k is

$$\mathcal{L}\{\xi_{\delta, k}(t)\} = \frac{2}{L\mu} \frac{1}{s^2 + (2/\tau_k')s + 1/\tau_k'^2 + 4\pi^2 f_k'^2}, \quad (\text{A5})$$

from which the low frequency response ($s \rightarrow 0$ for $f \ll f_k'$) of the resonator can be approximated as

$$\xi_k(t) \approx \frac{2L}{ESk^2\pi^2} F_{t \rightarrow l, k}(t), \quad (\text{A6})$$

which was obtained by writing $f_k' = k\sqrt{ES/\mu}/(2L)$ and assuming $1/\tau_k' \ll f_k'$. The shape of the dashed lines in Figs. 2(a) and (b) confirms that the longitudinal modes have constant gain at low frequencies. Note that this point of the derivation relates to neglecting the inertial effects, i.e., assuming $\mu(\partial^2 \xi / \partial x^2) = 0$ in Eq. (6). (See also Refs. 8 and 9.)

In this case, the component that comes from the longitudinal motion is expressed as

$$\begin{aligned} T_1(x, t) &= ES \frac{\pi}{L} \sum_{k=1}^{\infty} \xi_k(t) k \cos\left(\frac{k\pi x}{L}\right) \\ &= -\frac{2}{\pi} \sum_{k=1}^{\infty} \frac{1}{k} [F_{t \rightarrow l, k}(t)^+ \\ &\quad + F_{t \rightarrow l, k}(t)^-] \cos\left(\frac{k\pi x}{L}\right). \quad (\text{A7}) \end{aligned}$$

Calculating the excitation force $F_{t \rightarrow l, k}(t) = F_{t \rightarrow l, k}(t)^+ + F_{t \rightarrow l, k}(t)^-$ with the help of Eqs. (15a) and (15b) and eliminating k by substituting $m+n=k$ and $|m-n|=k$ gives

$$\begin{aligned} T_1(x, t) &= -ES \frac{\pi^2}{4L^2} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} y_m(t) y_n(t) mn \cos\left(\frac{m+n}{L} \pi x\right) \\ &\quad - ES \frac{\pi^2}{4L^2} \sum_{m=1}^{\infty} \sum_{\substack{n=1 \\ n \neq m}}^{\infty} y_m(t) y_n(t) mn \cos\left(\frac{m-n}{L} \pi x\right), \quad (\text{A8}) \end{aligned}$$

where $n \neq m$ in the second term comes from the fact that the longitudinal mode number $k = |m-n|$ cannot be zero. Note that there is no such constraint for the first term as $k = m+n$ in that case.

If Eqs. (A4) and (A8) are substituted into Eq. (A3), all the terms cancel out, except some with $m=n$ giving

$$T(x, t) = T_0 + ES \frac{\pi^2}{4L^2} \sum_{n=1}^{\infty} y_n^2(t) n^2 \cos\left(\frac{n-n}{L} \pi x\right), \quad (\text{A9})$$

which is the same as Eq. (A2) obtained by assuming uniform tension along the string. Note that the uniform tension approximation does not lead to zero longitudinal displacement. On the contrary, the string elements move in the longitudinal direction in a way that the tension remains uniform along the string. Actually, the tension could not be uniform without longitudinal motion.

The assumption leading to Eq. (A9) is that all the longitudinal modes are excited by frequencies that are considerably smaller than the corresponding longitudinal modal frequency f'_k . Thus, the validity of the uniform tension approximation should be evaluated by comparing f'_k with the excitation frequencies calculated by Eq. (16) for each k . Having small order transverse vibrations in comparison to the ratio of longitudinal and transverse propagation speeds (see, e.g., Ref. 8) is a sufficient, but not a necessary, condition for the applicability of the uniform tension approximation.

As a special case, if the transverse vibration contains only one mode, the uniform tension approximation can be applied. This is because the transverse mode n excites longitudinal mode $k = 2n$, for which Eq. (A6) holds. Note that this is true for all the transverse modes, and not only for the first few. However, when all the transverse partials are present up to a mode number N the assumption of the uniform tension can be applied only if the transverse vibration does not contain significant components around and above $f'_0/2$ (where $f'_0 = f'_1$ is the fundamental frequency of the longitudinal modal series). This is because the excitation force $F_{t \rightarrow 1,k}(t)$ has approximately double the bandwidth compared to the bandwidth of the transverse vibration.

It is important to note that the behavior of the longitudinal vibration changes qualitatively as a function of the transverse modes present on the string. It is an interesting field of future research to evaluate the coupling of transverse modes and the properties of nonplanar string vibrations in the case of realistic transverse components, i.e., when the uniform tension approximation cannot be applied. This would answer the question whether the theoretical and experimental results presented for the first few modes of transverse vibrations^{7–11} can be applied for the qualitative description of the phenomenon in the case of stringed instruments where dozens of transverse partials are generated.

¹H. A. Conklin, "Design and tone in the mechanoacoustic piano. Part III. Piano strings and scale design," *J. Acoust. Soc. Am.* **100**, 1286–1298 (1996).

²N. Giordano and A. J. Korty, "Motion of a piano string: Longitudinal vibrations and the role of the bridge," *J. Acoust. Soc. Am.* **100**, 3899–3908 (1996).

³I. Nakamura and D. Naganuma, "Characteristics of piano sound spectra," in *Proc. Stockholm Music Acoust. Conf.* (1993), pp. 325–330.

⁴H. A. Conklin, "Generation of partials due to nonlinear mixing in a stringed instrument," *J. Acoust. Soc. Am.* **105**, 536–545 (1999).

⁵J. Woodhouse, "Plucked guitar transients: Comparison of measurements and synthesis," *Acust. Acta Acust.* **90**, 945–965 (2004).

- ⁶B. Bank and L. Sujbert, "Modeling the longitudinal vibration of piano strings," in *Proc. Stockholm Music Acoust. Conf.*, Stockholm, Sweden (2003), pp. 143–146.
- ⁷R. Narasimha, "Nonlinear vibration of an elastic string," *J. Sound Vib.* **8**, 134–146 (1968).
- ⁸G. V. Anand, "Large-amplitude damped free vibration of a stretched string," *J. Acoust. Soc. Am.* **45**, 1089–1096 (1969).
- ⁹A. Watzky, "Non-linear three-dimensional large-amplitude damped free vibration of a stiff elastic stretched string," *J. Sound Vib.* **153**, 125–142 (1992).
- ¹⁰O. O'Reilly and P. J. Holmes, "Non-linear, non-planar and non-periodic vibrations of a string," *J. Sound Vib.* **153**, 413–435 (1992).
- ¹¹K. A. Legge and N. H. Fletcher, "Nonlinear generation of missing modes on a vibrating string," *J. Acoust. Soc. Am.* **76**, 5–12 (1984).
- ¹²T. Tolonen, V. Välimäki, and M. Karjalainen "Modeling of tension modulation nonlinearity in plucked strings," *IEEE Trans. Speech Audio Process.* **8**, 300–310 (2000).
- ¹³C. Erkut, M. Karjalainen, P. Huang, and V. Välimäki, "Acoustical analysis and model-based sound synthesis of the kantele," *J. Acoust. Soc. Am.* **112**, 1681–1691 (2002).
- ¹⁴J. O. Smith, "Physical modeling using digital waveguides," *Comput. Music J.* **16**, 74–91 (1992).
- ¹⁵S. Bilbao, "Energy-conserving finite difference schemes for tension-modulated strings," in *Proc. IEEE Int. Conf. Acoust., Speech, and Sign. Proc.*, Montreal, Canada (2004), pp. 285–288.
- ¹⁶M. J. Leamy and O. Gottlieb, "Internal resonances in whirling strings involving longitudinal dynamics and material non-linearities," *J. Sound Vib.* **236**, 683–703 (2000).
- ¹⁷E. V. Kurmyshev, "Transverse and longitudinal mode coupling in a free vibrating soft string," *Phys. Lett. A* **310**, 148–160 (2003).
- ¹⁸J. Bensa and L. Daudet, "Efficient modeling of phantom partials in piano tones," in *Proc. Int. Symp. on Musical Acoust.*, Nara, Japan (2004), pp. 207–210.
- ¹⁹B. Bank and L. Sujbert, "A piano model including longitudinal string vibrations," in *Proc. Conf. on Digital Audio Effects*, Naples, Italy (2004), pp. 89–94.
- ²⁰P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw–Hill, New York, 1968), Chap. 14.3, pp. 856–863.
- ²¹P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw–Hill, New York, 1968), Chap. 4.3, p. 120.
- ²²P. M. Morse, *Vibration and Sound* (American Institute of Physics, New York, 1983), Chap. III.10, p. 106, reprint of 2nd ed. (1st ed. 1936).
- ²³H. Fletcher, E. D. Blackham, and R. Stratton, "Quality of piano tones," *J. Acoust. Soc. Am.* **34**, 749–761 (1962).
- ²⁴G. Weinreich, "Coupled piano strings," *J. Acoust. Soc. Am.* **62**, 1474–1484 (1977).
- ²⁵Recorded and synthetic sound examples are available at <http://www.mit.bme.hu/~bank/publist/jasa-longitud>
- ²⁶N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments* (Springer-Verlag, New York, 1998), Chap. 12, pp. 352–398, 2nd ed. (1st ed. 1991).
- ²⁷A. Chaigne and A. Askenfelt, "Numerical simulations of piano strings. I. A physical model for a struck string using finite difference methods," *J. Acoust. Soc. Am.* **95**, 1112–1118 (1994).
- ²⁸A. Chaigne and A. Askenfelt, "Numerical simulations of piano strings. II. Comparisons with measurements and systematic exploration of some hammer-string parameters," *J. Acoust. Soc. Am.* **95**, 1631–1640 (1994).
- ²⁹B. Stopper, "Minimens 1.0 audio piano string simulator" (2003), URL: <http://www.piano-stopper.de/homepe.htm>
- ³⁰J. Bensa, S. Bilbao, R. Kronland-Martinet, and J. O. Smith, "The simulation of piano string vibration: From physical models to finite difference schemes and digital waveguides," *J. Acoust. Soc. Am.* **114**, 1095–1107 (2003).
- ³¹G. Bloothoof, E. Bringmann, M. van Cappellen, J. B. van Luipen, and K. P. Thomassen, "Acoustics and perception of overtone singing," *J. Acoust. Soc. Am.* **92**, 1827–1836 (1992).

Aerodynamic excitation and sound production of blown-closed free reeds without acoustic coupling: The example of the accordion reed

Denis Ricot^{a)}

Laboratoire de Mécanique des Fluides et d'Acoustique, UMR CNRS 5509, Ecole Centrale de Lyon, France

René Caussé and Nicolas Misdariis

Institut de Recherche et Coordination Acoustique/Musique, UMR CNRS 9912, 1 place Igor Stravinsky 75004 Paris, France

(Received 7 February 2003; revised 27 September 2004; accepted 2 December 2004)

The accordion reed is an example of a blown-closed free reed. Unlike most oscillating valves in wind musical instruments, self-sustained oscillations occur without acoustic coupling. Flow visualizations and measurements in water show that the flow can be supposed incompressible and potential. A model is developed and the solution is calculated in the time domain. The excitation force is found to be associated with the inertial load of the unsteady flow through the reed gaps. Inertial effect leads to velocity fluctuations in the reed opening and then to an unsteady Bernoulli force. A pressure component generated by the local reciprocal air movement around the reed is added to the modeled aerodynamic excitation pressure. Since the model is two-dimensional, only qualitative comparisons with air flow measurements are possible. The agreement between the simulated pressure waveforms and measured pressure in the very near-field of the reed is reasonable. In addition, an aeroacoustic model using the permeable Ffowcs Williams–Hawkings integral method is presented. The integral expressions of the far-field acoustic pressure are also computed in the time domain. In agreement with experimental data, the sound is found to be dominated by the dipolar source associated by the strong momentum fluctuations of the flow through the reed gaps. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1852546]

PACS numbers: 43.75.Pq [NHF]

Pages: 2279–2290

I. INTRODUCTION

The accordion is based on a very old system: the free metal reed which dates from the third millenium BC. It was used in a Chinese musical instrument, the tcheng. Nowadays, this system can be found in the harmonica and the harmonium, for example. The accordion reed is a thin metal plate riveted at one end to a support plate. There is a rectangular slot in the support plate immediately beneath the reed. The aperture is a bit larger than the reed so that the it freely vibrates as a cantilevered beam and does not strike against the support plate. An air flow is generated by the inward or outward movement of the bellows. With regards to Fletcher's classification,¹ the accordion reed is a valve of type $(-, +)$. In the presence of pre-existing flow in the same direction as the instrument, a steady overpressure applied from the upstream side of the reed tends to close the valve and inversely, a steady overpressure applied from downstream of the reed causes it to open further. Of course, if we consider a steady flow in the opposite direction, the valve is of type $(+, -)$. However, in the instrument, the accordion reed only operates in its blown-closed configuration: the inward movement of the bellows generates an upstream steady overpressure P_0 and the outward movement generates a depression on the downstream side of the reed which is equivalent to an upstream overpressure. The self-sustained oscillation results

from the unsteady pressure difference between both sides of the reed. But unlike most wind instruments such as woodwind and brass,^{2,3} the time-dependent pressure drop is not driven by the oscillation of the air column of a resonator. In the accordion, we can consider that the reed oscillation mechanism and the emitted sound do not depend on the acoustic influence of the upstream and downstream volumes. This kind of oscillating valve can be called a dominant or a strong reed.

In classical reed models^{1,2,4} that have been developed for resonator-coupled reeds, the pressure fluctuation $p_{ac}(t)$ is written from the acoustic response of the resonator excited by an unsteady volume flux $q_{ac}(t)$. By assuming the incompressibility of the flow around the reed, the acoustic volume flux at the resonator inlet is matched to the aerodynamic flux through the reed aperture $q_{ac}(t) := q(t) \propto V(t)x_2(t)L$ where $V(t)$ is the flow velocity, $x_2(t)$ is the reed displacement and L is a characteristic length of the orifice. The aerodynamic velocity—and thus the volume flux—is easily calculated from the steady-state Bernoulli equation with a pressure drop $P_0 - p_{ac}(t)$. This modeling approach implicitly supposes that the pressure of the reed flow at the exit of the reed channel exactly equals the acoustic pressure at the entrance of the resonator. This assumption can be justified⁵ by noting that a jet is formed at the exit of the reed channel because of the large abrupt transition in cross-sectional area from reed channel to instrument mouthpiece. All the kinetic energy of the flow is supposed to be dissipated in the spreading turbulent

^{a)}Current address: 10 avenue Paul Cezanne 78990 Elancourt, France; electronic mail: DenisRicot@aol.com

jet. This hypothesis is also consistent with the modeling of the acoustic source. Indeed, it is clear that only the fluctuating mass is taken for the excitation of the acoustic air column. The fluctuating force associated with the periodic momentum injection can be ignored. Finally, the reed is modeled by a simple mass-spring-damping system. Consequently, a coupled system of three equations (the impedance relation, the Bernoulli equation and the simple harmonic oscillator equation) and three unknown variables [$x_2(t), q(t), p_{ac}(t)$] can be written. With the same approach, the influence of the internal acoustic impedance of the air supply system (the player's lungs, vocal tract and mouth) can be taken into account.¹ Because of the impedance relation, most calculations are performed in the frequency domain.^{1,2} But the frequency-domain approach is not advantageous for strongly nonlinear equations such as the Bernoulli equation. Time-domain solutions^{6,4} and hybrid methods⁷ have also been proposed.

For a constant surface of contact between the flow and the reed, the dynamic excitation force is a combination of acoustic pressure and pressure induced by velocity variation due to the variation in the aperture area. The velocity-induced pressure fluctuations are found by applying the stationary Bernoulli equation which associates a decrease of the pressure with an acceleration of the fluid. But, without other physical mechanisms such as acoustic feedback, these pressure fluctuations cannot induce self-oscillations. For example, if we suppose a constant volume flow through the reed aperture, it is straightforward to show that the force induced by fluid velocity variation is in phase with the displacement. Energy transfer between the constant supply pressure and the reed oscillation is not possible if we consider only the stationary Bernoulli equation. Titze⁸ explained that an asymmetry of the pressure force during the cycle is necessary for reed excitation. In case of resonator-coupled reeds, the pressure asymmetry is provided by the acoustic response of upstream and/or downstream volumes. In case of the accordion reed, there is no acoustic feedback. By supposing that the stationary Bernoulli equation well describes the aerodynamic flow through the reed aperture, other physical mechanisms have been proposed to explain the phase shift of the total force with regard to the reed displacement. For example, a complex evolution of the channel geometry during the cycle can lead to the asymmetry of the total force. This approach is used to model voice production because vocal fold oscillations are not considered to be driven by the acoustic impedance of lungs and vocal tract. Titze⁸ used a wave motion model to describe the movement of the tissues and Pelorson *et al.*⁹ added a second mechanical oscillator: a phase delay of the total pressure force is obtained because the shape variations of the channel are not the same during closing and opening of the glottis.

The effect of separation and reattachment of the flow in the reed channel has also been studied. Indeed, the flow separation at the inlet of the clarinet reed channel leads to a local acceleration of the flow⁵ (vena-contracta effect). Moreover, the total pressure force can also depend on the possible reattachment of the air flow on the reed further in the reed channel. The vena-contracta factor and the location of the reat-

tachment point may vary as a function of the reed position and produce a force in phase with the reed velocity. An example of the modeling of these viscous effects can be found in the study of Pelorson *et al.*⁹ At the outlet of the vocal fold channel, he studied the movement of the point where the air flow separates from the glottis. In this case, the total pressure force fluctuates because the reed surface in contact with the unsteady flow fluctuates. Although the consequences of the flow separation are qualitatively and quantitatively significant, it was not shown whether this phenomenon could induce self-oscillation without the presence of the two-mass model.

The separation points of the accordion reed flow are fixed at the sharp edge of the reed. As shown in Sec. II of this paper, there is no flow reattachment on the other side of the reed, therefore the surface of the reed immersed in the flow is constant. The flow separation/reattachment effects are not a candidate for a possible excitation mechanism of the accordion reed.

If we discard the supposition that the unsteady flow can be described by the stationary Bernoulli equation, it appears that the energy transfer from the flow to a free reed can be obtained by two other physical mechanisms. First, the vortex shedding behind a solid obstacle¹⁰ induces a periodic force on the structure. Thanks to their experimental investigations, Saint Hilaire *et al.*¹¹ concluded that vortex shedding is not responsible for the excitation of his harmonium reed. This result has been also confirmed in a recent study¹² for a blown-open reed.

At last, the inertial effect of the upstream flow has been proposed to be the excitation mechanism of free reeds¹¹ in the absence of acoustic coupling. The verification of this assumption is one of the purposes of the present work. An analysis of the flow around the accordion reed (Sec. II) shows that the flow can be modeled as an unsteady potential two-dimensional flow. In Sec. III, the aerodynamic force which excites the vibration of the reed is calculated under the incompressible flow hypothesis. From aerodynamic data, a model for the calculation of the aeroacoustic source based on an acoustic analogy is proposed (Sec. IV). Flow and aeroacoustic models are solved in the time domain and results are compared with experimental data (Sec. V). Since the attention of this paper is turned to the analysis of the exact nature of both the excitation force and acoustic source, empirical hypotheses or adjustable parameters must not be used. Because of the complexity of flow equations, the calculation of analytical expressions is only possible with a very simplified representation of the accordion reed. Therefore quantitative values of physical variables are not expected to agree with the actual ones. Only waveforms and spectra are compared to measurements.

II. FLUID MECHANICAL DESCRIPTION OF THE FLOW

The geometry of the accordion reed is slightly different from the reed geometry used by Saint Hilaire *et al.*¹¹ in their flow visualizations. The initial distance x_0 between the vibrating reed and the support plate is smaller for the accordion reed than for the Saint Hilaire arrangement. The reed is directly riveted onto the support plate, therefore x_0 is very

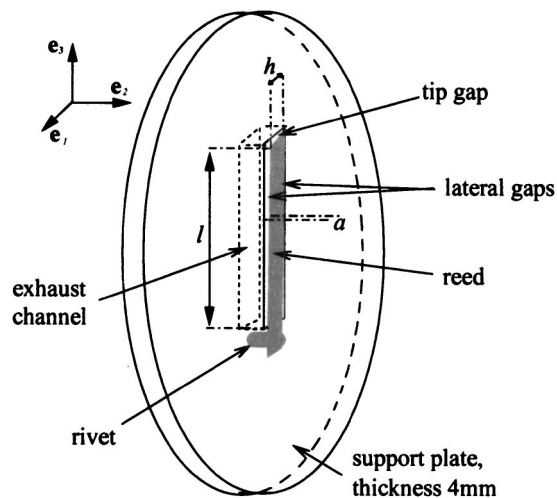


FIG. 1. Sketch of the Plexiglas support plate and riveted reed. The flow is from right to left. The mathematical axes used in the model (Secs. III and IV) are also shown.

small and greatly depends on the stationary profile of the reed. In our case, we can suppose that $x_0 \leq 0.4$ mm. The thickness of the support plate is another difference. The slot in the support plate will be called the exhaust channel in the following (see Fig. 1). The length of the exhaust channel in the streamwise direction is equal to the thickness of the support plate. We see that this length is quite large compared to the width of the gaps. Therefore a flow reattachment is expected to occur on the lateral walls of the exhaust channel.

In order to check the validity of Saint Hilaire's conclusions concerning the inability of wake instabilities to excite the reed and to have more details on the flow structure, flow visualizations using dye stream injections have been conducted in water. The experimental arrangement is given in Fig. 2. The accordion reed is mounted on a Plexiglas reservoir of about 800 cm³. The steel reed was $l=32$ mm in length and $b=0.2$ mm in thickness and has a mean width of $h=3.6$ mm. In air, the natural frequency of the reed is $f_0=330$ Hz. The parameter a is the remaining distance between the reed and the exhaust channel walls when the reed

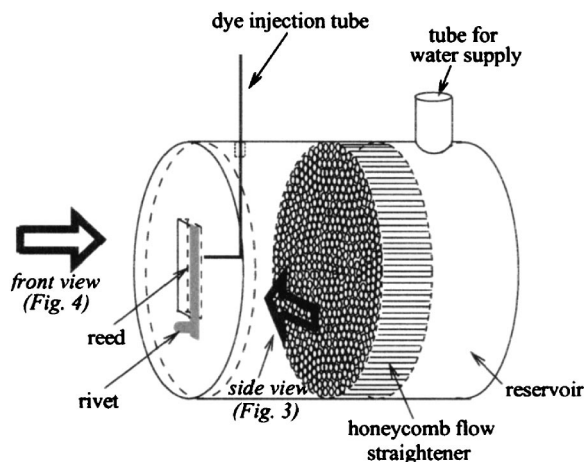


FIG. 2. The experimental setup for water-flow visualizations. This setup is also used for measurements of reed displacement, aerodynamic and acoustic pressures in air flow (Sec. V). The two large arrows indicate the views of the photographs shown in Fig. 3 and Fig. 4.

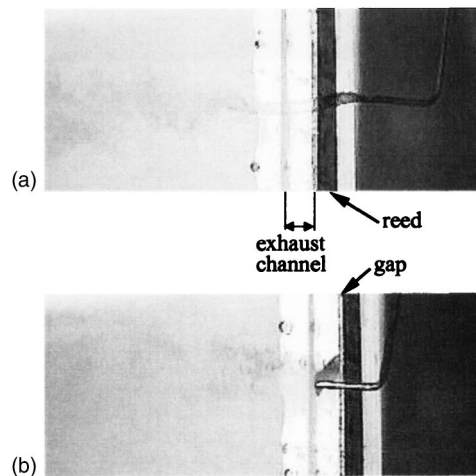


FIG. 3. Side views (see Fig. 2 and Fig. 5) of the flow pattern for two dye stream injection points situated (a) in front of a reed gap, and (b) in the support plate plane. Only a central part of the reed is shown. The supply cavity is on the right. In the upstream region, the dye is convected along potential streamlines. Downstream of the reed, turbulent mixing and diffusion are visible.

is closed. It means that the width of the exhaust channel is $h+2a$. Generally, a is less than 0.2 mm. The notations for the reed dimensions are given in Fig. 1. The reed mounted on its supply cavity is immersed in a discharge tank. The height of the free surface of water in the discharge tank is maintained constant thanks to an overflow pipe. The water was supplied to the reed cavity through a narrow flexible tube from a large supply tank. The average supply pressure of the reed (the difference between the upstream and the downstream pressures) was simply set by varying the distance between the free surfaces of water of the supply and discharge tanks. A honeycomb flow straightener ensures an even upstream flow. This system is visible in the background of the photography given later in Fig. 4. Visualization of the flow field was accomplished by the injection, via a capillary tube, of a dye stream at various locations upstream of the reed.

For a supply pressure of $P_0=2000$ Pa, self-sustained oscillations of the reed can be observed. The mean velocity \bar{V} of the flow through the reed gaps can be evaluated using the stationary Bernoulli equation. The Reynolds number based on the reed width h is about $Re_h = \bar{V}h/\nu \approx 6000$ where ν is the kinematic viscosity of water. For self-sustained oscillations in air at normal playing pressure, the Reynolds number of the flow is slightly inferior to that observed in water but has the same order of magnitude. Thus, we can suppose that the excitation mechanism and the flow characteristics in water are the same as that in air flow. The frequency of reed oscillation under water flow is 142 Hz. The stroboscopic images that are not reported here clearly show that the reed oscillates on the first mode of the cantilever beam. The great difference between the oscillation frequency and the natural frequency of the reed in air is a consequence of the added mass effect of water (see Sec. III C).

Typical flow patterns are shown in Figs. 3(a) and 3(b). As expected, the upstream flow is laminar. The dye is convected along well-defined streamlines. The dye streams show

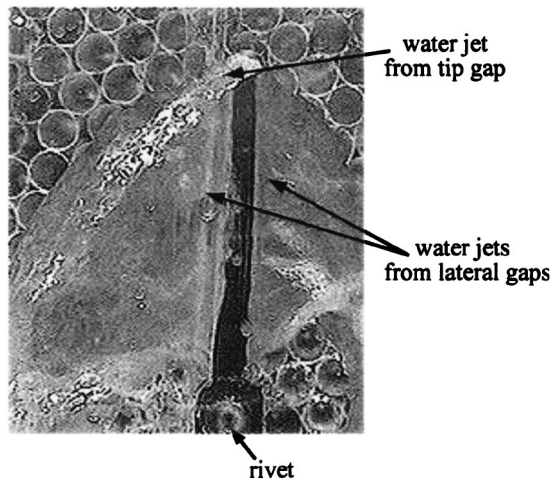


FIG. 4. Front view (see Fig. 2 and Fig. 5) of the water flow around the reed if the water discharge tank is empty. The reed also oscillates in this configuration. Water plane-jets are formed from the lateral and tip gaps. The downstream side of the reed (that is visible in the picture) is in a dry region.

that it is consistent to model the upstream flow by potential sinks. Moreover, for injection points far enough from the tip gap, streamlines lie in planes $x_3 = \text{constant}$, where x_i is the coordinate along the axis \mathbf{e}_i (see Fig. 1). Thus, the flow upstream of the reed seems to be two-dimensional. This assumption is used in the proposed flow model to describe the near-field potential flow. The outflow is turbulent. In photographs, we see the mixing and the diffusion of the dye jets downstream of the reed. However, the characteristic sizes of the turbulent structures are small compared to the reed width h . There are no large coherent vortices that could have been responsible for the reed excitation. Because of the very small gap sizes compared to the reed width and to the thickness of the support plate, the downstream flow does not have the characteristics of a wake. In fact, it appears that the jet is attached to the lateral wall of the exhaust channel.

Figure 4 shows the water flow around the reed if the discharge tank is empty so that the downstream fluid is air. Self-oscillations also occur in this configuration. The added mass effect is reduced in comparison with the first configuration because only one side of the reed is immersed in water. Thus, it is not surprising to find a vibration frequency of 206 Hz that is superior to that obtained when the upstream and downstream fluid is water. The flow view of Fig. 4 shows that the downstream flow can be considered as plane jets, which do not interact with the downstream side of the reed. Even though this configuration is unrealistic, it is interesting because it definitively shows that the downstream flow does not take part in the reed oscillation mechanism.

III. AERODYNAMIC EXCITATION AND REED OSCILLATION

A. Dynamics of the potential flow

The flow visualizations in water show that only the upstream face of the vibrating plate is in contact with the moving fluid. Therefore the excitation force must be evaluated by integrating the aerodynamic pressure fluctuation over the upstream face. The flow through the reed gaps can be supposed

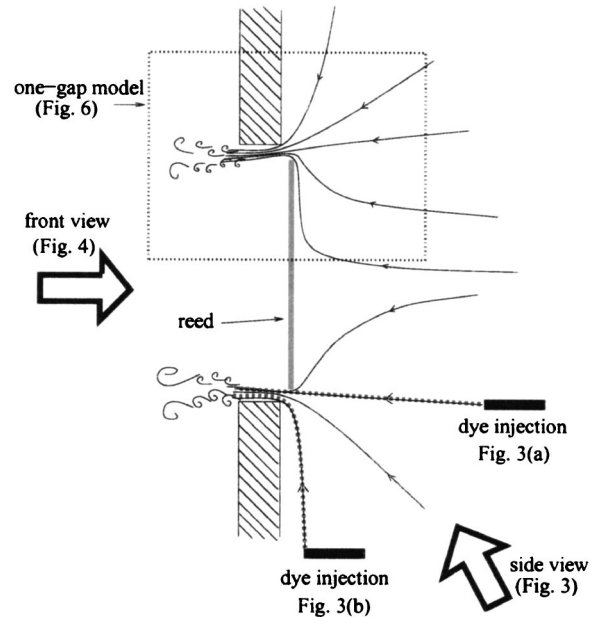


FIG. 5. Sketch of the flow patterns through the reed gaps in a plane $x_3 = \text{constant}$. The upstream flow is a potential flow induced by the sinks (the lateral gaps). Downstream of each gap, the flow is a plane wall-jet, and it turns to a turbulent free-jet after separation from the exhaust channel walls.

incompressible because the Mach and Helmholtz numbers are very small. Even if the downstream jets are turbulent at the exit of the exhaust channel, it can be considered to be laminar and irrotational to a certain extent from the gaps. Consequently, potential expressions are used to model the upstream and downstream flows. In this analysis, the inviscid two-dimensional flow is described in a section perpendicular to the length of the reed (plane $x_3 = \text{constant}$). This section is supposed to be far enough from the tip of the reed to avoid three-dimensional effects. Using the superposition property of potential flows,¹³ the total flow could be found by the summation of the identical potential flows created by the two lateral gaps drawn in Fig. 5. However, in order to obtain the most simple mathematical expressions, it is reasonable to suppose that the physics of reed excitation can be described considering simply one-gap flow. This very simplified representation of a blown-closed free reed is shown in Fig. 6.

The width of the gap is $e(t) = \sqrt{x_2^2(t) + a^2}$ if $x_2(t) > 0$ and $e(t) = a$ if $x_2(t) < 0$. The flow past the reed is approximated by assuming that the reed lies in the support plate plane $x_2 = 0$. The potential of the upstream flow is expressed as the sum of a continuous distribution of fluid sinks,¹³

$$\phi_1(x_1, x_2, t) = -\frac{1}{\pi} V(t) \int_0^{e(t)} \ln \left(\frac{\sqrt{(x_1 - u)^2 + x_2^2}}{e(t)/2} \right) du, \quad (1)$$

where $V(t)$ is the velocity of the downstream jet. The jet velocity $\mathbf{V}_2 = -V(t)\mathbf{e}_2$ is supposed to be constant across the gap, along the axis \mathbf{e}_1 . The integral of Eq. (1) can be analytically evaluated. The expressions of $\phi_1(x_1, x_2, t)$ and the velocity $\mathbf{V}_1(x_1, x_2, t) = \mathbf{grad}(\phi_1)$ are given in Appendix A. The logarithmic singularity in the expressions of the potential and velocity at points (0,0) and $(e(t), 0)$ are a classical consequence of the inviscid potential model. At these points, the

C. Reed motion and reciprocal flow

It has been previously shown that the motion of a free reed^{14,15} is almost perfectly sinusoidal. Laser vibrometer system measurements that are not reported here have also been made on our accordion reeds and the results are consistent with the previous observations. Thus the reed can be modeled by a mass-spring-damping single oscillator system. The reed motion leads to a reciprocal fluid current that is not dependent on the main unsteady flow. The contribution of the reed surface vibration to the total pressure can be studied using the radiation impedance of a baffled rectangular piston $Z_m(\omega) = F_2(\omega)/u(\omega)$, where $u(\omega)$ is the normal velocity of the reed and $F_2(\omega)$ is the reaction force of the fluid back on the driving piston for an angular frequency ω . For $\omega l/c_0 \ll 1$, where c_0 is the speed of sound, the largest term in Z_m is its imaginary part, the reactance. Morse and Ingard¹⁶ calculated the expression of the reactance of a baffled rectangular piston for one of its faces,

$$F_2(\omega) = -i\omega m_f u(\omega), \quad \text{with}$$

$$m_f = \frac{8}{9\pi} \rho_0 l h \frac{l^2 + lh + h^2}{l+h}. \quad (9)$$

This force F_2 is in an opposite phase with the reed acceleration and it describes that the fluid entrained by one reed side has an apparent mass of m_f . Consequently, the effective fluid mass $2m_f$ must be added to the accelerated reed mass m in the two-dimensional reed motion equation,

$$\frac{(m+2m_f)}{l} \frac{\partial^2 x_2}{\partial t^2} + \frac{c}{l} \frac{\partial x_2}{\partial t} + \frac{k}{l} (x_2 - x_0) = F_1(t), \quad (10)$$

where c and k are the equivalent damping and stiffness of the reed. The added mass causes a decrease of the natural frequency of the reed. In air, this effect is very small but in a heavier fluid such as water the difference between the intrinsic natural frequency and the measured natural frequency can be very large. The reed oscillator coefficients have been determined by a classical experimental procedure based on static stiffness evaluation and free decay behavior. These measurements are not detailed in this paper. The oscillator mass is determined from the measurement of the natural frequency of the reed in air. The experimental effective mass then includes both the intrinsic mass m and the added mass of air $2m_f$. In fact, the study of the reed-induced reciprocal flow is only useful for the estimation of the near-field resultant pressure that is added to the aerodynamic pressure $P_1(t)$ for comparisons with experimental data (Sec. V). In the near field of the upstream region, the mean fluctuating pressure associated with the reed reactance is

$$P_2(t) = -\frac{1}{lh} m_f \frac{\partial^2 x_2}{\partial t^2}. \quad (11)$$

Moreover, it is worth noting that the reciprocal flow associated with the reed reactance must not be taken into account for the calculation of the far-field acoustic pressure.

In this section, we have written in the time domain the equations of flow and reed motion. If we notice that $V(t)$ and $\partial V/\partial t$ can be expressed as a function of $x_2(t)$ and $q(t)$ and

their time derivative, we see that the model of reed excitation is a differential system of two coupled nonlinear equations (5) and (10) with the unknowns $x_2(t)$ and $q(t)$. A fourth order Runge–Kutta algorithm is used for the numerical system solving. The nonconstant coefficients of the differential system are updated at each step of the Runge–Kutta procedure.

IV. AEROACOUSTIC MODEL

As for other wind instruments, the sound produced by the accordion is a consequence of the unsteady flow through the reed apertures. But unlike resonator-coupled instruments, we cannot deduce the sound from the analysis of the resonator response to the inlet unsteady flux. For the accordion reed, we can consider that the fluctuating fluid flow acts directly as a source of sound on the free acoustic medium. Therefore we neglect in the model all contributions of reflection or diffraction of acoustic waves on the neighboring solid walls.

A. Lighthill's acoustic analogy

The problem of sound production by flow was first investigated by Lighthill¹⁷ in 1952. His theory is based on the comparison between the exact equations of fluid motion with the equations of sound propagation in a medium at rest. The mass and momentum conservation equations are rewritten to form the most general inhomogeneous wave equation,

$$\frac{\partial^2 \rho'}{\partial t^2} - c_0^2 \nabla^2 \rho' = \frac{\partial^2 T_{ij}(\mathbf{y}, t)}{\partial y_i \partial y_j} \quad (12)$$

where $\rho' = \rho - \rho_0$ and $p' = p - P_{\text{atm}}$. The tensor $T_{ij} = \rho u_i u_j + (p' - c_0^2 \rho') \delta_{ij} - \tau_{ij}$ is the Lighthill stress tensor where δ_{ij} is the Kronecker delta and τ_{ij} is the viscous stress tensor. The summation convention on indices i and j is used. For high Reynolds, low Mach number flows and in the absence of thermal sources the Lighthill stress tensor can be reduced¹⁷ to $T_{ij} \approx \rho_0 u_i u_j$. With this simplification, the source term of Eq. (12) expresses the variation of the rate of momentum flux which is induced by turbulent processes. The upstream sink flow is not turbulent and, arguing that the upstream and downstream acoustic waves are the same but with an opposite sign, it becomes evident that the turbulent sources are not important in the generation of an accordion reed sound for both the blowing and drawing notes. However, the physics of aeroacoustic sources can only be analyzed by using the mathematical expressions of the far acoustic field. The far-field acoustic pressure is calculated using the convolution product of Lighthill's source with the free-space Green's function. In this case, the spatial derivatives of Lighthill's tensor with respect to the source position must be rewritten in terms of derivatives with respect to the observer's position. By a careful mathematical analysis^{18,19} of the bounded volume integration of Lighthill's tensor, mass and momentum sources appear as residual terms of the Lighthill tensor flux through the surface that bounds the volume. These monopole and dipole terms on the surface add to the quadrupole source generated by turbulence in the control volume. This analysis was first proposed by Curle,¹⁸ and

Ffowcs Williams and Hawkings²⁰ gave a rigorous mathematical framework in order to take into account solid or fictitious (permeable) control surfaces and moving solid bodies.

B. Application of the permeable Ffowcs Williams–Hawkings integral method

The Ffowcs Williams–Hawkings (FW–H) integral method²⁰ was first used for the calculation of noise generated both by turbulence and moving impenetrable surfaces, but recently the FW–H equation has been validated as an integral formulation with permeable surfaces.^{21,22}

The derivation of the FW–H formulation is based on the mathematical technique of generalized functions. A generalized function is formed with the aid of Heavisides’s function $H(f)$ defined to be unity where $f > 0$ and zero where $f < 0$. The equation $f = 0$ defines the control surface Σ .

The Navier–Stokes equations are written for the generalized density $H(f)\rho'$, momentum $H(f)\rho u_i$, and pressure $H(f)p'$. As for Lighthill’s equation (12), an inhomogeneous wave equation can be expressed from the mass and momentum conservation equations. Details of the derivation of the FW–H equation from conservation laws are given by Brentner and Farassat.²³ For a nonmoving permeable surface, the inhomogeneous wave equation is

$$\begin{aligned} & \frac{\partial^2 \rho' H(f)}{\partial t^2} - c_0^2 \nabla^2 \rho' H(f) \\ &= \frac{\partial^2 T_{ij} H(f)}{\partial y_i \partial y_j} - \frac{\partial}{\partial y_i} \left((\rho u_i u_j + p \delta_{ij} - \tau_{ij}) \delta(f) \frac{\partial f}{\partial y_j} \right) \\ &+ \frac{\partial}{\partial t} \left(\rho u_j \delta(f) \frac{\partial f}{\partial y_j} \right). \end{aligned} \quad (13)$$

Generally, the control surface Σ is supposed to enclose all the volume sources described by Lighthill’s tensor T_{ij} . Then, $T_{ij}H(f) = 0$ for $f \geq 0$. Moreover, as the free-field Green’s function is used for the calculation of the far-field acoustic pressure, all the possible flow/acoustic interactions and acoustic wave reflections on solid walls must be taken into account as flow variables on the control surface. Because of our incompressible approach, all these effects are neglected in the source terms. Finally, neglecting the viscous stress tensor in the second term of the right hand side of the FW–H equation, we see that the acoustic sources can be divided into a monopole Q and a dipole F_i that are calculated on the permeable control surface,

$$\begin{aligned} Q(\mathbf{y}, t) &= \rho u_j \frac{\partial f}{\partial y_j} \approx \rho_0 u_j \frac{\partial f}{\partial y_j}, \\ F_i(\mathbf{y}, t) &= -(\rho u_i u_j + p \delta_{ij}) \frac{\partial f}{\partial y_j} \approx -(\rho_0 u_i u_j + p \delta_{ij}) \frac{\partial f}{\partial y_j}. \end{aligned} \quad (14)$$

The free-field solution of Eq. (13) is the convolution product of the source terms with the free-space Green’s function. After rearrangement of the integral expressions (see Appendix C), the density fluctuations are given by

$$\begin{aligned} \rho'(\mathbf{x}, t) &= \frac{1}{4\pi c_0^2} \frac{\partial}{\partial x_i} \int_{\Sigma} \left[\frac{\rho_0 u_i u_j + p \delta_{ij}}{r} \right] n_j d\Sigma \\ &+ \frac{1}{4\pi c_0^2} \frac{\partial}{\partial t} \int_{\Sigma} \left[\frac{\rho_0 u_j}{r} \right] n_j d\Sigma, \end{aligned} \quad (16)$$

where the square brackets denote that the functions are evaluated at the retarded time $\tau = t - r/c_0$ and n_j are the components of the unit outward normal vector to surface. When $|\mathbf{x}| \ll L \ll \lambda$, where L is a typical dimension of the control surface Σ and λ is a typical wavelength of the sound generated, a classical simplification procedure^{18,17} leads to the final expression of the acoustic pressure,

$$\begin{aligned} p_{ac}(\mathbf{x}, t + |\mathbf{x}|/c_0) &= \frac{1}{4\pi c_0} \frac{x_i}{|\mathbf{x}|^2} \frac{\partial}{\partial t} \int_{\Sigma} (\rho_0 u_i u_j + p \delta_{ij}) n_j d\Sigma \\ &+ \frac{\rho_0}{4\pi |\mathbf{x}|} \frac{\partial}{\partial t} \int_{\Sigma} u_j n_j d\Sigma. \end{aligned} \quad (17)$$

Obviously, the second integral on the right hand side of Eq. (17) is the expression of the total volume flux through the control surface that is evidently equal to $q(t)$. The first source term of Eq. (17) has a dipolar form and its value is evaluated numerically using an adaptive recursive Simpson’s rule.

V. EXPERIMENTS AND MODEL VALIDATION

A. Experimental study in air

The experimental setup for air measurements is similar to that used for water visualizations (Fig. 2). The water supply tank is replaced by a low-impedance pressure source that insures a constant blowing pressure inside the reservoir. A probe microphone measures the aerodynamic pressure at a few millimeters from the upstream face of the reed. A variable capacitance transducer is used for the motion of the reed. This transducer is based on a condenser microphone, the membrane of which is electrically replaced by the reed. A water manometer measures the average supply pressure and a microphone captures the radiated sound. This microphone is located at 25 cm downstream of the reed.

In Fig. 7, displacement, aerodynamic pressure and acoustic pressure are represented for a reed with a natural frequency of $f_0 = 795$ Hz. The characteristic dimensions of the valve are $l = 22.9$ mm, $h = 2.5$ mm, $a = 0.14$ mm, and $x_0 = 0.2$ mm. The supply pressure is $P_0 = 40$ Pa. As expected, we see in Fig. 7(a) that the reed motion is composed of sinusoidal oscillations around a displaced equilibrium position. We note that for low playing pressure, the steady displacement of the reed is small and we can consider that the equilibrium position is about x_0 . Figure 7(b) shows a sharp aerodynamic pressure spike at the moment when the reed enters the slot. Since the equilibrium position is not so far from the plane $x_2 = 0$, the force induced by this strong pressure variation is nearly in phase with the reed velocity. Thus, it provides energy to the oscillator. Similarly at the reed opening, a negative pressure variation occurs but it is less violent. The downstream acoustic signal shown in Fig. 7(c)

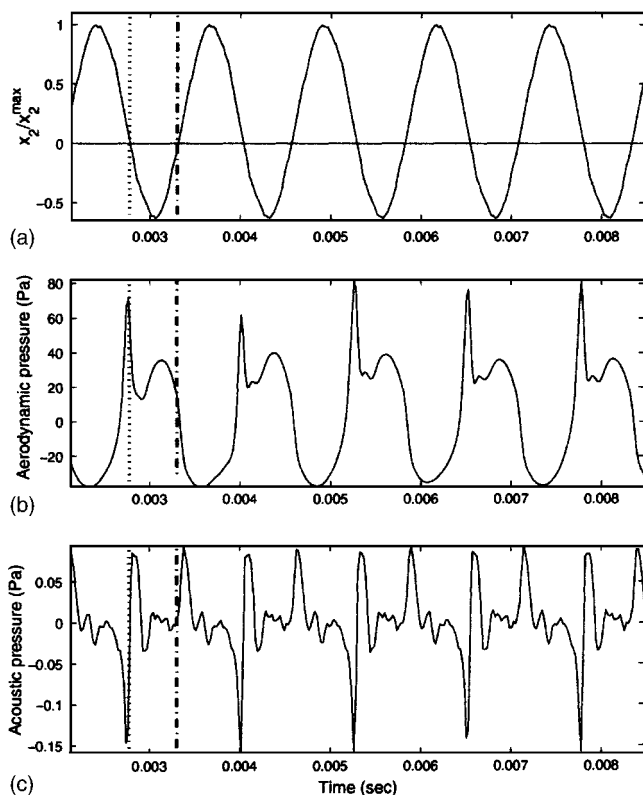


FIG. 7. Experimental measurements in air flow: (a) reed displacement; (b) upstream aerodynamic pressure; (c) downstream acoustic pressure. \cdots , closing of the valve; $-\cdots-$, opening of the valve.

greatly differs from the aerodynamic pressure waveform. This illustrates the complex conversion mechanism of the fluid mechanical energy into acoustic energy. Only the sharp pressure spike associated with the reed closure is common to the two waveforms with an opposite sign. This reverse sign results from the quasi-symmetry of the upstream and downstream flow with regard to the aperture plane. Figure 8 shows the spectrum of the acoustic pressure. In the rich spectral distribution of the reed sound, the odd harmonics are predominant. It is interesting to note that both the waveform and

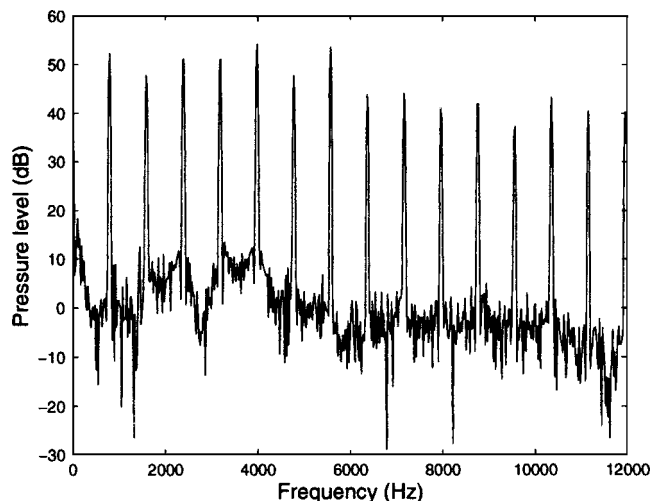


FIG. 8. Spectrum of the measured acoustic pressure.

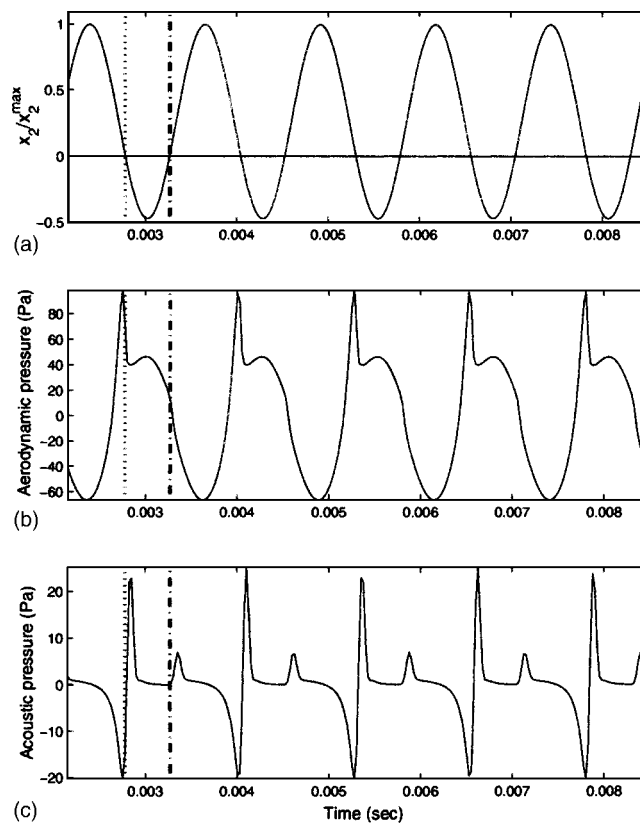


FIG. 9. Results of the model calculation: (a) reed displacement; (b) aerodynamic pressure at point C; (c) reverse of the acoustic waveform at point D. \cdots , closing of the valve; $-\cdots-$, opening of the valve.

spectrum of the acoustic signal compare well with signals extracted from recordings of accordion notes.²⁴

B. Application of the model

For the calculation, the model variables are initialized at zero and the reed position is set to x_0 . The differential equation system, Eqs. (5) and (10), is solved using Eq. (8) for the excitation force. After the transient, self-sustained sinusoidal oscillations are observed at the natural frequency of the reed [see Fig. 9(a)]. Using the computed fluid variables $q(t)$ and $V(t)$, the aerodynamic pressure at point C of Fig. 6 is computed from the Bernoulli equation (3) with the complete expression of the potential derivative, Eq. (A3). For a comparison with the probe microphone signal, the contribution $P_2(t)$ of the reciprocal flow is added to this aerodynamic excitation pressure. The total signal is plotted in Fig. 9(b). The agreement with the measured waveform is good.

Since the downstream flow model is very simplified compared to the actual development of turbulent spreading jets, the acoustic pressure is calculated in the upstream region. The symmetry property of the produced sound is used to compare the calculated signal to the acoustic measurement performed in the downstream region. The modeled acoustic pressure is the external acoustic pressure that would be measured if the reed oscillated in its sucking configuration. The fluid variables calculated on the control surface Σ defined in Fig. 6 are introduced in the integral equation (17). The result is plotted in Fig. 9(c). Even if the pressure spike occurring just after the reed opening is smaller than the measured one,

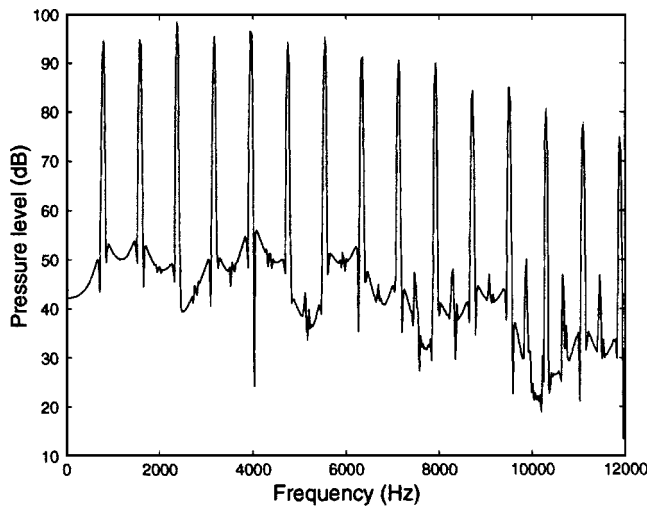


FIG. 10. Spectrum of the modeled acoustic pressure.

the modeled waveform reproduces quite well the experimental acoustic signal. Moreover, it is shown in Fig. 10 that the odd harmonics of the modeled sound spectrum are more powerful than the even ones, which is a characteristic of the accordion reed sound. Examples of synthesized sounds are available on the IRCAM website.²⁵

C. Analysis of the aerodynamic excitation

From the previous comparisons, we can suppose that the physics of both reed excitation and acoustic emission are well reproduced by the model. The calculation results can be used to highlight some details of the physical mechanisms.

For example, Fig. 11 shows the components of the aerodynamic pressure force $F_1(t)$ associated with the Bernoulli force [the first term of Eq. (8)] and inertial force [second and third terms of Eq. (8)]. This figure clearly demonstrates that the inertial force is a negligible part of the total force. In most reed models,^{1,5,2,4} inertial forces are neglected. This simplification can also be done for the accordion reed model and from Eq. (8), the average excitation pressure on the upstream face of the reed is now

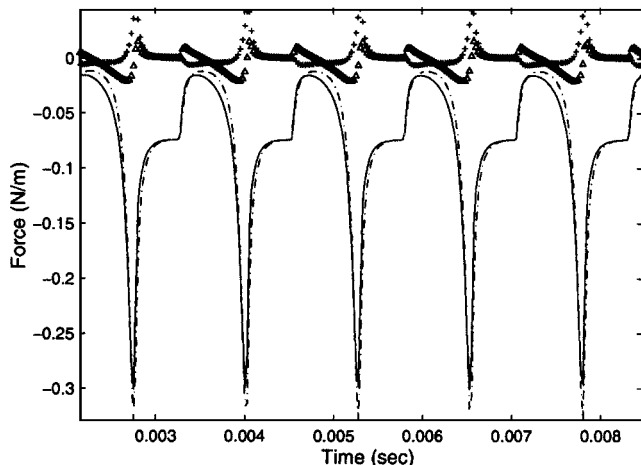


FIG. 11. A comparison of the three terms of the aerodynamic force given by Eq. (8): ---, first term; Δ , second term; +, third term; -, total force.

$$\langle P_1(t) \rangle = -\frac{1}{h} F_1(t) = \frac{\rho_0}{2} V^2(t) (1 - A_1(t)). \quad (18)$$

Replacing the fluid velocity in Eq. (18) by its expression given by Eq. (5), the average pressure becomes

$$\langle P_1(t) \rangle = -\rho_0 K_0(t) (1 - A_1(t)) \frac{\partial q}{\partial t} + P_0 (1 - A_1(t)). \quad (19)$$

As explained before, the contraction effect associated with $A_1(t)$ does not contribute to the excitation mechanism. Thus, the term $P_0(1 - A_1(t))$ on the right hand side of Eq. (19) is not an excitation pressure. Finally, the expression of the excitation pressure is

$$\langle P_1(t) \rangle_{exc} = -\rho_0 K_0(t) (1 - A_1(t)) \frac{\partial q}{\partial t}. \quad (20)$$

This relation looks like an impedance relation giving the pressure as a function of the volume flux. In this case, the impedance is inertial. The value of the equivalent ‘‘acoustic’’ inertia $M(t) = \rho_0 K_0(t) (1 - A_1(t))$ is time-dependent.

Unlike traditional theoretical approaches of oscillating valves¹ based on acoustic impedances of upstream and downstream volumes, Eq. (20) is related to the impedance of the orifice. $M(t)$ can be interpreted as the mass end correction¹⁶ of the upstream side of the gap. The mass end correction depends on the detailed form of the potential flow and gap geometry. Thus, the time dependence of $M(t)$ is due to the variation of gap size. In particular, $K_0(t)$ is related to the total quantity of fluid that takes part in the inertial mechanism. For the calculation presented in this paper, the average value of the parameter K_0 is 4.5. This corresponds to a very large distance of the reference point B where the pressure is supposed to be constant and the fluid velocity is zero. The choice of point B is arbitrary but first applications of the aerodynamic model show that the influence of K_0 on the result becomes small above a certain value that depends on dimensions and characteristics of the simulated reed.

D. Analysis of the sound production

In addition, the presented model allows us to analyze the mechanism of sound production. The monopole and dipole sources of Eq. (17) are represented in Fig. 12. It appears immediately that the sound of the reed is dominated by the dipolar component. This observation is in opposition to the physics of the reed coupled to a resonator. Indeed, as mentioned in the Introduction, the acoustic source exciting the air column of the resonator is modeled as the variation of the volume flux through the reed gaps. In the case of a free reed, the monopolar source associated with unsteady volume flux at the orifice is negligible, the acoustic waves are excited by the violent changes of momentum flux through the orifice.

VI. CONCLUSIONS

In this paper, a simplified representation of a blown-closed reed used, for example, in accordions has been studied. The flow around the reed is described using an incompressible potential flow theory and the excitation force in this

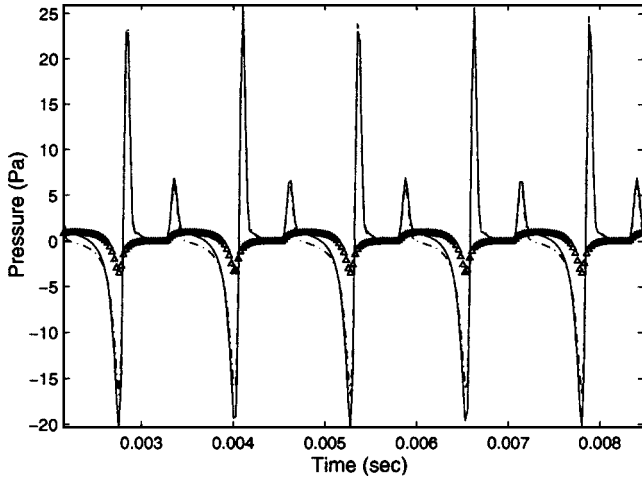


FIG. 12. A comparison of the monopole and dipole terms of the acoustic expression, Eq. (17): Δ , monopole term; $-\cdot-$, dipole term; $-$, total aeroacoustic source.

model is associated with the inertial load of the upstream fluid. The delay in the variation of upstream fluid velocity due to the inertia leads to a velocity fluctuation in the reed opening. Unlike for resonator-coupled reeds, the pressure-velocity relationship is not driven by the impedance of upstream and downstream volumes but by an equivalent inertial impedance of the gaps. One of the difficulties arises from the time-dependence of the orifice geometry which leads to a time-dependent mass inertia.

This work also provides the first model of the source of sound generated by a free reed in the absence of a resonator. Despite the numerous simplifications made in the calculation, the qualitative agreement of the synthesized pressure signal with the measured one in terms of waveform and spectrum is quite good. The model is based on the application of an acoustic analogy that allows the derivation of the acoustic pressure in the far-field as a function of the aerodynamic variables known in the source region. In our case, the acoustic source is calculated in the time domain but it is also possible to perform the calculation in the frequency domain.

ACKNOWLEDGMENTS

The authors wish to thank Agnès Maurel and Philippe Petitjeans of the Ecole Supérieure de Physique et de Chimie Industrielles for helpful discussions about the flow visualization arrangement. We also want to thank Alain Terrier and Gérard Bertrand of IRCAM for their help in the measurements.

APPENDIX A: CALCULATION OF THE UPSTREAM FLOW

In this appendix, the analytical expressions of the flow potential and velocity for the upstream region are given.

1. Flow potential in the upstream region

The upstream potential is given by Eq. (1):

$$\phi_1(x_1, x_2, t) = -\frac{1}{\pi} V(t) \int_0^{e(t)} \ln \left(\frac{\sqrt{(x_1 - u)^2 + x_2^2}}{e(t)/2} \right) du. \quad (\text{A1})$$

The coordinates x_1 and x_2 are given in Fig. 6. $V(t)$ is the fluid velocity through the gap and $e(t)$ is the gap size. After calculations, one can obtain the explicit expression of the potential:

$$\begin{aligned} \phi_1(x_1, x_2, t) = & -\frac{1}{\pi} V(t) \left(\frac{1}{2} e(t) \ln \left(\frac{(x_1 - e(t))^2 + x_2^2}{(e(t)/2)^2} \right) \right. \\ & - \frac{1}{2} x_1 \ln \left(\frac{(x_1 - e(t))^2 + x_2^2}{x_1^2 + x_2^2} \right) + x_2 \left(\arctan \left(\frac{x_1}{x_2} \right) \right. \\ & \left. \left. - \arctan \left(\frac{x_1 - e(t)}{x_2} \right) \right) \right) + \frac{1}{\pi} q(t), \quad (\text{A2}) \end{aligned}$$

where $q(t) = V(t)e(t)$ is the volume flux through the gap.

2. Time derivative of the flow potential in the upstream region

The time derivative of the flow potential given by Eq. (A2) is

$$\begin{aligned} \frac{\partial \phi_1(x_1, x_2, t)}{\partial t} = & \frac{1}{\pi} \frac{\partial V}{\partial t} \left(\frac{1}{2} x_1 \ln \left(\frac{(x_1 - e(t))^2 + x_2^2}{x_1^2 + x_2^2} \right) \right. \\ & \left. - x_2 \left(\arctan \left(\frac{x_1}{x_2} \right) - \arctan \left(\frac{x_1 - e(t)}{x_2} \right) \right) \right) \\ & + \frac{1}{\pi} \frac{\partial q}{\partial t} \left(1 - \ln \left(\frac{\sqrt{(x_1 - e(t))^2 + x_2^2}}{e(t)/2} \right) \right). \quad (\text{A3}) \end{aligned}$$

3. Components of the velocity vector in the upstream region

The two components of the upstream velocity vector are found by calculating the gradient of the potential ϕ_1 [Eq. (A2)]:

$$\begin{aligned} V_{1,x_1}(x_1, x_2, t) = & \frac{\partial \phi_1}{\partial x_1} = \frac{V(t)}{2\pi} \ln \left(\frac{(x_1 - e(t))^2 + x_2^2}{x_1^2 + x_2^2} \right), \quad (\text{A4}) \\ V_{1,x_2}(x_1, x_2, t) = & \frac{\partial \phi_1}{\partial x_2} \\ = & \frac{V(t)}{\pi} \left(\arctan \left(\frac{x_1 - e(t)}{x_2} \right) - \arctan \left(\frac{x_1}{x_2} \right) \right). \quad (\text{A5}) \end{aligned}$$

APPENDIX B: EXPRESSION OF THE AERODYNAMIC FORCE

The total force on the upstream face of the reed is given by integration of the aerodynamic pressure [Eq. (6)] over the reed width h ,

$$\mathbf{F}_1(t) = - \left(\int_{e(t)}^{e(t)+h} P(x_1, 0, t) dx_1 \right) \mathbf{e}_2. \quad (\text{B1})$$

The force then has the form

$$F_1(t) = -h \frac{\rho_0}{2} V^2(t) (1 - A_1(t)) + \rho_0 h^2 A_2(t) \frac{\partial V}{\partial t} - \rho_0 h A_3(t) \frac{\partial q}{\partial t}, \quad (\text{B2})$$

with the parameters

$$A_1(t) = \frac{1}{h \pi^2} \int_{e(t)}^{e(t)+h} \left(\ln \left(\frac{x_1 - e(t)}{x_1} \right) \right)^2 dx_1, \quad (\text{B3})$$

$$A_2(t) = \frac{1}{h^2 \pi} \int_{e(t)}^{e(t)+h} x_1 \ln \left(\frac{x_1 - e(t)}{x_1} \right) dx_1, \quad (\text{B4})$$

$$A_3(t) = \frac{1}{h \pi} \int_{e(t)}^{e(t)+h} \ln \left(\frac{x_1 - e(t)}{e(t)/2} \right) dx_1. \quad (\text{B5})$$

For an analytical calculation of the integrals, we can note that the logarithmic singularities in potential and velocity expressions are integrable.

The final expressions of the parameters A_1 , A_2 , and A_3 are given below.

1. Parameter $A_1(t)$

$$A_1(t) = \frac{1}{\pi^2} \left(\ln^2(h) + \left(1 + \frac{e(t)}{h} \right) \ln(e(t)+h) \right) \times \ln \left(\frac{e(t)+h}{h^2} \right) + \frac{e(t)}{h} \ln(e(t)) \ln \left(\frac{e(t)}{h^2} \right) - 2 \frac{e(t)}{h} \operatorname{dilog} \left(\frac{e(t)+h}{e(t)} \right), \quad (\text{B6})$$

where the dilogarithm function is defined by

$$\operatorname{dilog}(x) = \int_1^x \frac{\ln(u)}{1-u} du. \quad (\text{B7})$$

2. Parameter $A_2(t)$

$$A_2(t) = \frac{1}{2\pi} \left(\ln(h) + \frac{e(t)}{h} (\ln(h^2) - 1) + \frac{e^2(t)}{h^2} \ln(e(t)) - \frac{(e(t)+h)^2}{h^2} \ln(e(t)+h) \right). \quad (\text{B8})$$

3. Parameter $A_3(t)$

$$A_3(t) = \frac{1}{\pi} \left(\ln \left(\frac{2h}{e(t)} \right) - 1 \right). \quad (\text{B9})$$

APPENDIX C: CALCULATION OF THE ACOUSTIC DENSITY FLUCTUATIONS

The free-field Green's function is

$$G(\mathbf{x}, t | \mathbf{y}, \tau) = \frac{1}{4\pi c_0^2 r} \delta(g), \quad \text{with } g = \tau - t + r/c_0, \quad r = \|\mathbf{x} - \mathbf{y}\|. \quad (\text{C1})$$

where \mathbf{y} is a point in the source region and \mathbf{x} is the observer's position. Then the acoustic density fluctuations are

$$\rho'(\mathbf{x}, t) H(f) = \frac{1}{4\pi c_0^2} \int_{\mathcal{R}^3} \int_{\mathcal{R}} \frac{1}{r} \left(\frac{\partial}{\partial y_i} F_i(\mathbf{y}, t) \delta(f) \right) \delta(g) d\boldsymbol{\tau} d\mathbf{y} + \frac{1}{4\pi c_0^2} \int_{\mathcal{R}^3} \int_{\mathcal{R}} \frac{1}{r} \frac{\partial}{\partial t} (Q(\mathbf{y}, t) \delta(f)) \delta(g) d\boldsymbol{\tau} d\mathbf{y}. \quad (\text{C2})$$

The spatial integral of the first term can be developed as

$$\int_{\mathcal{R}^3} \frac{1}{r} \left(\frac{\partial}{\partial y_i} F_i(\mathbf{y}, t) \delta(f) \right) \delta(g) d\mathbf{y} = \int_{\mathcal{R}^3} \frac{\partial}{\partial y_i} \left(\frac{1}{r} F_i(\mathbf{y}, t) \delta(f) \delta(g) \right) d\mathbf{y} - \int_{\mathcal{R}^3} F_i(\mathbf{y}, t) \delta(f) \frac{\partial}{\partial y_i} \left(\frac{1}{r} \delta(g) \right) d\mathbf{y}. \quad (\text{C3})$$

The first integral of the right hand side is zero and noting that $\partial(\delta(g)/r)/\partial y_i = \partial(\delta(g)/r)/\partial x_i$ we can write

$$\int_{\mathcal{R}^3} \frac{1}{r} \left(\frac{\partial}{\partial y_i} F_i(\mathbf{y}, t) \delta(f) \right) \delta(g) d\mathbf{y} = - \frac{\partial}{\partial x_i} \int_{\mathcal{R}^3} \frac{1}{r} F_i(\mathbf{y}, t) \delta(f) \delta(g) d\mathbf{y} = \frac{\partial}{\partial x_i} \int_{\Sigma} \frac{1}{r} (\rho_0 u_i u_j + p \delta_{ij}) n_j \delta(g) d\Sigma, \quad (\text{C4})$$

with

$$n_j = \frac{1}{|\nabla f|} \frac{\partial f}{\partial y_j}. \quad (\text{C5})$$

The second spatial integral of the density fluctuations can be also written as

$$\int_{\mathcal{R}^3} \frac{1}{r} \frac{\partial}{\partial t} (Q(\mathbf{y}, t) \delta(f)) \delta(g) d\mathbf{y} = \frac{\partial}{\partial t} \int_{\Sigma} \frac{1}{r} \rho_0 u_j \delta(g) n_j d\Sigma. \quad (\text{C6})$$

Finally, the temporal integration of the two transformed expressions of the spatial integrals immediately gives the formulation of Eq. (16).

¹N. H. Fletcher, "Autonomous vibration of simple pressure-controlled valves in gas flows," *J. Acoust. Soc. Am.* **93**, 2172–2180 (1993).

²J. Saneyoshi, H. Teramura, and S. Yoshikawa, "Feedback oscillations in reed woodwind and brasswind instruments," *Acustica* **62**, 194–210 (1987).

³S. C. Thompson, "The effect of the reed resonance on woodwind tone production," *J. Acoust. Soc. Am.* **66**, 1299–1307 (1979).

⁴R. T. Schumacher, "Ab initio calculations of the oscillations of a clarinet," *Acustica* **48**, 71–85 (1981).

⁵A. Hirschberg, R. W. A. van de Laar, J. P. Marrou-Maurières, A. P. J. Wijnands, H. J. Dane, S. G. Kruijswijk, and A. J. M. Houtsma, "A quasi-stationary model of air flow in the reed channel of a single-reed woodwind instruments," *Acustica* **70**, 146–154 (1990).

⁶B. Gazengel, J. Gilbert, and N. Amir, "Time domain simulation of single reed wind instrument. From the measured input impedance to the synthesis signal. Where are the traps?," *Acta Acust. (Beijing)* **3**, 445–472 (1995).

⁷J. Gilbert, J. Kergomard, and E. Ngoya, "Calculation of the steady-state

- oscillations of a clarinet using the harmonic balance technique," *J. Acoust. Soc. Am.* **86**, 35–41 (1989).
- ⁸I. R. Titze, "The physics of small-amplitude oscillation of the vocal folds," *J. Acoust. Soc. Am.* **83**, 1536–1552 (1988).
- ⁹X. Pelorson, A. Hirschberg, R. R. van Hassel, and A. P. J. Wijnands, "Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. Application to a modified two-mass model," *J. Acoust. Soc. Am.* **96**, 3416–3431 (1994).
- ¹⁰R. D. Blevins, *Flow-Induced Vibration* (Krieger, Malabar, Florida, 1986).
- ¹¹A. O. St. Hilaire, T. A. Wilson, and G. S. Beavers, "Aerodynamic excitation of the harmonium reed," *J. Fluid Mech.* **49**, 803–816 (1971).
- ¹²A. Z. Tarnopolsky, J. C. S. Lai, and N. H. Fletcher, "Flow structures generated by pressure-controlled self-oscillating reed valves," *J. Sound Vib.* **247**, 213–226 (2001).
- ¹³A. R. Paterson, *A First Course in Fluid Dynamics* (Cambridge University Press, Cambridge, 1983), pp. 205–242.
- ¹⁴J. P. Cottingham, C. J. Lilly, and C. H. Reed, "The motion of air-driven free reeds," *Proceedings of the 2nd Convention of the European Acoustics Association: Forum Acusticum*, Berlin, 1999.
- ¹⁵L. Millot, C. Cuesta, and C. Valette, "Experimental results when playing chromatically on a diatonic harmonica," *Acustica* **87**, 262–270 (2001).
- ¹⁶P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).
- ¹⁷M. J. Lighthill, "On sound generated aerodynamically. I. General theory," *Proc. R. Soc. London, Ser. A* **221**, 564–587 (1952).
- ¹⁸N. Curle, "The influence of solid boundaries upon aerodynamic sound," *Proc. R. Soc. London, Ser. A* **321**, 505–514 (1955).
- ¹⁹A. P. Dowling and J. E. Ffowcs Williams, *Sound and Sources of Sound* (H Ellis Horwood, Chichester, 1983), pp. 157–166.
- ²⁰J. E. Ffowcs Williams and D. L. Hawkings, "Sound generation by turbulence and surfaces in arbitrary motion," *Philos. Trans. R. Soc. London, Ser. A* **264**, 321–342 (1969).
- ²¹P. di Francescantonio, "A new boundary integral formulation for the prediction of sound radiation," *J. Sound Vib.* **202**, 491–509 (1997).
- ²²X. Gloerfelt, C. Bailly, and D. Juvé, "Computation of the noise radiated by a subsonic cavity using direct simulation and acoustic analogy," *AIAA Paper No. 2001-2226*, 2001.
- ²³K. S. Brentner and F. Farassat, "An analytical comparison of the acoustic analogy and Kirchhoff formulation for moving surface," *AIAA J.* **36**, 1379–1386 (1998).
- ²⁴N. Misdariis, D. Ricot, and R. Caussé, "Modélisation physique de la vibration d'une anche d'accordéon," *Proceedings of the 5th French Congress on Acoustics*, Lausanne, 2000, pp. 281–283.
- ²⁵<http://www.ircam.fr/equipes/instruments/> website: last time viewed by the authors: November 2004.

Measurement of the depth-dependent resonance of water-loaded human lungs

J. S. Martin and P. H. Rogers

School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0405

E. A. Cudahy

Naval Submarine Medical Research Laboratory, Groton, Connecticut 06349-5900

(Received 16 June 2004; revised 17 December 2004; accepted 22 December 2004)

An experiment was conducted to determine the response of the human lung to water-borne sound in the range of 20 to 500 Hz. A small pool inside a hyperbaric chamber was used to simulate four ambient pressure conditions spanning the range of recreational diving depths. Ten subjects were tested on two occasions each using three separate measures to evaluate the response of the subjects' lungs. With some notable exceptions, results were consistent between subjects and between measures. These indicate that human lungs can be reasonably modeled as a lumped single-degree-of-freedom system over the lower portion of the band of interest. Here, the surrounding fluid provides the dominant mass and the dominant stiffness is provided by the entrapped air with a small additional contribution from tissue elasticity. Measured resonances increase with the square root of ambient pressure from an average of 40 Hz with a quality factor of 1.8 at near-surface pressure to 73 Hz with a quality factor of 2.6 at an equivalent depth of 36.4 m. There is evidence of other resonances within or near the band of interest that may be attributable to nonvolumetric chest/lung modes, Helmholtz resonance, and/or resonance of gastrointestinal bubbles. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1858311]

PACS numbers: 43.80.Cs, 43.80.Ev, 43.30.Jx [FD]

Pages: 2291–2300

I. BACKGROUND

Advances in low-frequency active (LFA) sonar systems have raised concerns for the safety of divers that may be exposed to the acoustic radiation produced by these systems.^{1–3} Similar concerns have also been raised for the safety of sea mammals.^{1,4–6} Compliance with safe exposure levels for divers and sea mammals can dictate operational parameters for LFA sonar systems beyond the limits imposed by the physics of transduction and acoustic propagation and are therefore of great concern to system operators.⁷ Unfortunately, damage and aversion thresholds for both divers and marine life are not well understood. Physical damage thresholds are based on evidence that is anecdotal or on experiments with small animal models.⁶ Aversion thresholds have been derived from experiments that were conducted under conditions that make an extrapolation of the results to real-world scenarios difficult.^{1,2} To match existing data to these scenarios, it is necessary to develop models for the response of organ systems to incident acoustic signals in the free field. The most significant of these systems, from a predictive standpoint are those that resonate within the frequency range of a particular LFA system such as the US Navy's Surveillance Towed Array Sonar System (SURTASS) LFA. These can either be the most likely organs to be damaged by the incident sound or can act as secondary sources that mediate the exposure level of other organ systems. The SURTASS LFA frequency range is sufficiently low that the organs most likely to resonate are compliant bodies such as the lungs and other gas-filled spaces. The experiment reported here was intended to test the response of gas-filled spaces within the human body that might resonate in the SURTASS LFA fre-

quency range. Although this experiment does not provide underwater-sound-exposure criteria, the information may be of use in the interpretation of exposure criteria for depths other than those that were used in threshold-determination experiments.

II. CONFIGURATION OF THE EXPERIMENT

Ideally, the response of the human body to low-frequency underwater sound should be tested under free-field conditions. These can be simulated at sea, in a very large lake or tank, or in a water tank with effectively anechoic boundaries. The frequencies of interest for the SURTASS LFA problem make either of the latter two schemes impractical if only passive boundaries are considered. These boundaries would need to attenuate signals with wavelengths of up to 75 m at the lowest frequencies of interest. At-sea testing was deemed to be impractical because of cost and logistical constraints, and because of the potential danger to the test subjects. As an alternative to these schemes a configuration was sought in which the nature of the expected response could be used to reduce the scale of the experiment. Because the lungs are much smaller than a wavelength in water at even the highest frequencies of interest ($f = 500$ Hz, $\lambda = 3$ m) it is reasonable to assume that the radiation load imposed on their dilatational motion by the water will be dominantly reactive (i.e., mass-like). Similarly, the load imposed on a uniformly vibrating body within a small tank with soft boundaries will also be mass-like because the water within the tank moves in unison with the surface of the vibrating body, and the radiation resistance must be nearly zero because very little energy can be radiated across the bound-

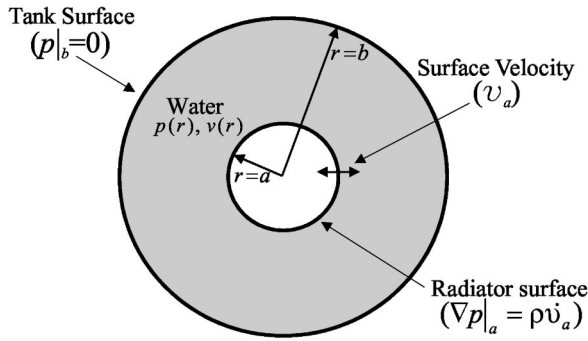


FIG. 1. Spherically symmetric encapsulated radiator model for the impedance loading imposed by a small water tank.

aries. Given that the radiation impedance on the lung would be mass-like both in the small tank and in a free field, it is reasonable to explore the circumstances under which these two impedances would be effectively identical. The analysis of a simple, related system sheds considerable light on this problem.

If a sphere of radius a oscillates radially at the center of a water-filled spherical enclosure with radius b and a pressure-release outer boundary (as depicted in Fig. 1), then the spherically symmetric Helmholtz equation can be used to describe the acoustic field within the water:

$$\nabla^2 P + k^2 P = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial P}{\partial r} \right) + k^2 P = 0. \quad (1)$$

Here $P(r)$ is the spatially dependent complex amplitude of the acoustic pressure and k is the wave number in the fluid. Two boundary conditions apply to this equation. The first of these is the pressure-release condition at the outer boundary and the second is a prescribed radial acceleration at the surface of the inner sphere:

$$P|_{r=b} = 0, \quad \nabla P|_{r=a} = -i\omega\rho v_a. \quad (2)$$

Here v_a is the radial velocity of the surface, ρ is the quiescent density of the fluid, and Euler's equation relates the pressure gradient to acceleration. The differential equation and boundary conditions can be satisfied by a spherical standing wave of the form

$$P_{\text{encl}} = \frac{-i\omega\rho v_a a^2}{ka \cos(k(b-a)) + ka \sin(k(b-a))} \times \left[\frac{\sin(kb)\cos(kr) - \cos(kb)\sin(kr)}{r} \right]. \quad (3)$$

In the absence of the outer boundary the Sommerfeld radiation condition would apply and the exterior field would be described by an outward traveling spherical wave of the form

$$P_{\text{ff}} = \left(\frac{-i\omega\rho v_a a^2}{e^{ika}[1 - ika]} \right) \left(\frac{e^{ikr}}{r} \right). \quad (4)$$

The impedance measured at the surface of the radiator ($r = a$) in either of the cases defined in Eqs. (3) and (4) completely defines the fluid loading on the radiator. Taking the ratio of the impedance at the surface of the encapsulated

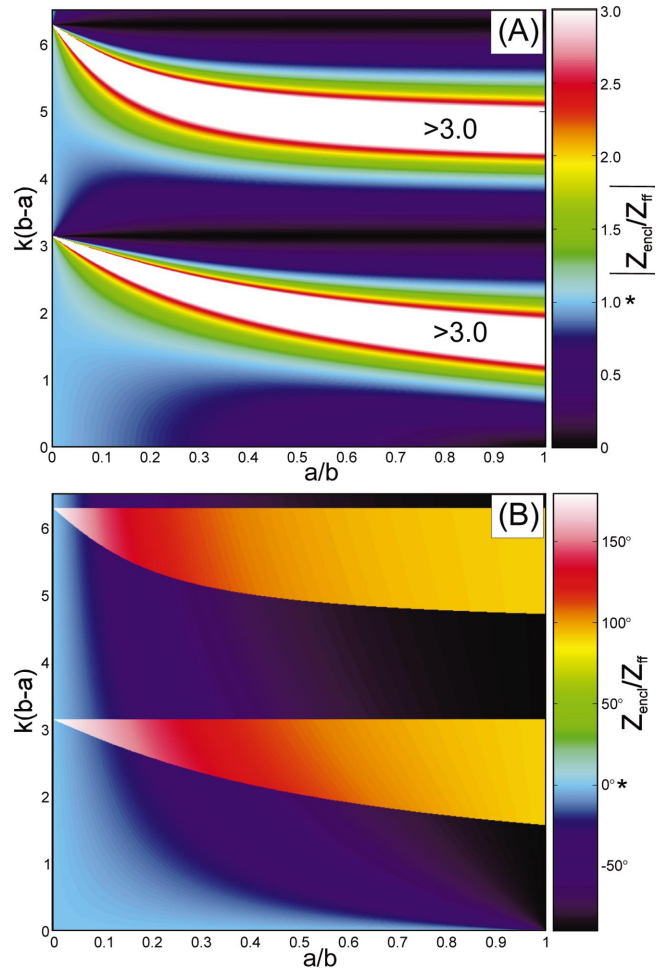


FIG. 2. Impedance imposed by a tank, computed from the encapsulated sphere model, relative to free field: amplitude (a) and phase (b).

sphere to the surface impedance of the sphere in free field gives the expression

$$\frac{Z_{\text{encl}}}{Z_{\text{ff}}} = \frac{\left[\frac{P_{\text{encl}} \nabla P_{\text{ff}}}{P_{\text{ff}} \nabla P_{\text{encl}}} \right]_{r=a}}{ka \cot(k(b-a)) + 1} = \frac{1 - ika}{ka \cot(k(b-a)) + 1}, \quad (5)$$

where $Z = (P/v)_{r=a}$ for the subscripted scenario.

It is clear from this expression that the impedance of the encapsulated sphere can only approximate the free-field impedance when ka is small so that the numerator is dominantly real and the corresponding free-field impedance is reactive. Physically, this is because the fluid has no mechanism for extracting energy from the encapsulated sphere, but it can radiate away in the free field. It is also clear from Eq. (5) that the cotangent of $k(b-a)$ must be small because the impedance is very far from correct when $ka \cot(k(b-a))$ either has a large magnitude or is close to -1 . Physically, this is equivalent to the statement that resonances of the fluid volume must be avoided. These occur for values of $k(b-a) = n\pi$ for $n=0,1,2,\dots$, and for $kb \sim n\pi$ if ka is assumed to be very small. These conditions can be easily avoided if $kb < \pi/2$ and $a < b$, where the latter is an obvious geometrical requirement. The impedance ratio can be plotted in terms of the nondimensional variables $k(b-a)$ and a/b , as shown in Figs. 2(a) and 2(b). Here, a third physical intuition is reinforced. This is that the impedance is only correct if the encapsulated sphere is much smaller than its enclosure, thus

ensuring that most of the kinetic energy of the free near-field would be contained within the enclosure. Elementary considerations show that, in a free field, the total nonpropagating kinetic energy of the fluid in a spherical shell from radius a out to radius r is given by

$$T(r) = \frac{2\pi\rho a^3 v_a^2 (1 - a/r)}{(1 + k^2 a^2)} \quad (6)$$

Equation (6) approaches the radiation mass of a free-field sphere (which is three times the displaced mass of water if $ka \ll 1$) when $r \rightarrow \infty$. From Eq. (6) it is apparent that 90% of the radiation mass is contributed by fluid within 10 radii of the center of the sphere.

Human lungs have a volume on the order of 4 l. The available pool, which was designed for other hyperbaric-chamber experiments, held approximately 4000 l. Thus a/b for the experiment was about 1/10th ($\sqrt[3]{1/1000}$). The equivalent spherical radii of the lung and pool volumes are about 10 cm and 1 m respectively. Thus, $0.1 < k(b-a) < 2$, so the experiment was conducted in a region depicted in the lower left of Figs. 2(a) and 2(b), where the impedance condition on the lung-like sphere is correct to about 10% in amplitude and 10° in phase at worst. At the low end of the frequency range, where lung resonances were observed, the impedance is much closer to the free-field condition. For small gas bodies such as bubbles in the stomach and intestines, the impedance in the tank is more nearly correct than it is for the lungs. The tank was designed to fit into the hatch of the hyperbaric chamber that housed it rather than being based on the above-mentioned acoustic considerations. In spite of this, it was very nearly ideal for the lung-resonance experiment. Because the lung volume and frequency range were fixed, a significantly larger or smaller tank would have placed the experiment in a region of either a higher $k(b-a)$ or a/b value where impedance conditions are less favorable. The tank was constructed with eight sections of 6.4 mm thick aluminum that were bolted together with soft rubber gaskets. The tank was open at the top and sat inside of the much larger hyperbaric chamber. There was only contact between the tank and the chamber where the bottom of the tank rested on foam tiles on the chamber floor. The interior surfaces of the tank had been covered with a 2.5 cm layer of closed-cell foam. There were no apparent in-band structural resonances of the tank walls that would have violated the assumptions of the spherical impedance model. The combination of the tank's free surface, thin-walled construction, and the surface treatment reasonably mimicked pressure release assumption of the spherical model. The tank was irregularly shaped, but its first mode was not strongly dependent on this geometry and could be predicted by a spherical model. This mode was measured at 950 Hz, an octave above the band of interest for the experiment. More detailed tank models have been developed for measurements at or near tank resonances.⁸ However a more detailed model of the tank was unnecessary for this experiment given that the wavelengths were long in comparison to it and that a detailed model of the lungs was unavailable to embed within such a model. It should be noted that detailed lung models have been developed for other

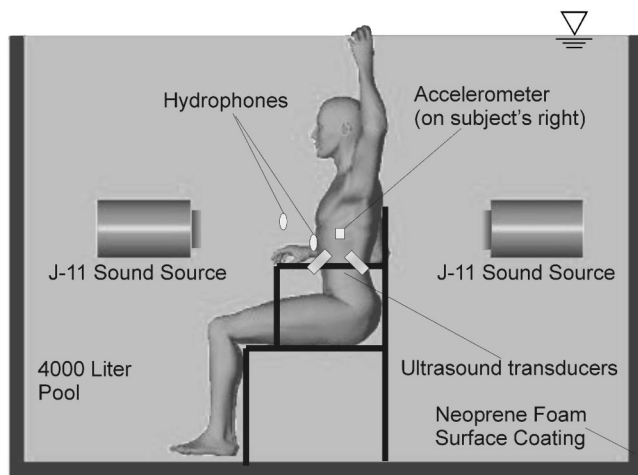


FIG. 3. Setup of an experimental water tank inside the hyperbaric chamber (chamber not shown).

applications.^{9,10} Unfortunately, these consider excitations and loading conditions inappropriate to the problem under consideration here.

III. EXPERIMENTAL METHODS

For the experiments, each diver was seated with his upper torso in the approximate center of the tank. USRD J-11 electrodynamic transducers¹¹ were suspended in front of and behind the diver at the mid-depth along the longest dimension of the tank. This configuration is shown in Fig. 3. These two transducers were driven with identical signals so that there was coherent summation of the pressure signals and destructive interference of the velocity at the tank's center. In a free-field environment and in the farfield of a source, a subject would be exposed to an incident signal with the acoustic pressure and velocity related by the intrinsic impedance of the fluid ($P/|\vec{v}| = \rho c$). It is not feasible to achieve this condition over a significant volume in a small tank and it is not obvious which orientation of the velocity vector would be appropriate for a worst-case scenario. It is, however, likely that the acoustic pressure rather than the velocity controls lung motion in the free field. This has been shown to be a reasonable assumption in the analysis of the low-frequency response of air bubbles and other compliant scatterers.^{12,13}

Ten subjects were tested in this experiment. All of the subjects were navy divers in good health, ranging from 26 years to 41 years in age and weighing between 64 and 98 kg. There were nine male subjects and one female. Subjects were selected for their availability, willingness to participate, and ability to participate in the experimental protocol.¹⁴ Pulmonary function was tested on all of the subjects. These tests included both vital capacity and lung volume measures. The mean full vital capacity for the subjects was 6.4 l and varied by $\pm 30\%$ over the test group with a weak correlation to the subject's body mass. Residual lung volumes averaged 1 l with a $\pm 55\%$ variability and no obvious correlation to weight. Air was supplied to the subjects by a diving regulator connected to an air supply outside the tank. During the tests, the subjects were instructed to inhale a volume of air with which they were comfortable holding their breath for ~ 20 s

and that permitted them to remain stationary in the focus of an ultrasonic measurement system. Because of this, it is likely that individual preferences overwhelmed physiological differences in determining the lung volumes during the measurements. Each subject was tested twice. Each test was performed at ambient pressures (measured in the hyperbaric chamber, i.e., at the free surface of the water tank) equivalent to 0, 2.1, 17.4, and 35.7 m of seawater (0, 7, 57, and 117 ft \pm 1 ft). The subjects were seated such that their chests were approximately centered at the mid-depth of the tank, which adds an additional 0.7 m to the equivalent depth because this was used as the datum. There was some variation in this because of the differing heights of the subjects. The depths and time at each depth were determined by decompression requirements. The time constraints imposed by decompression were such that higher pressures and other equivalent depths were not practical for these tests.

Sound exposure levels were not considered as a variable of interest in this study. Incident levels were set sufficiently low to ensure the linearity of the lung response and to meet the safe exposure guideline that was in place at the time of the study, which was 160 dB (*re* 1 μ Pa rms) based on pure-tone exposure at any frequency in the range and at any depth. Levels were set sufficiently high to ensure an adequate SNR within the time available for each measurement. Each subject was exposed to a sound pressure level of 150 dB (*re* 1 μ Pa rms) computed over the 16 s duration of the broadband exposure signal. No subjects reported discomfort from this exposure.

Four techniques were used to monitor the response of the subjects' lungs during the experiments. These are depicted in Fig. 3. Two of the measurements were made with B&K™ 8103 miniature hydrophones. One of these was in contact with the subject's abdomen just below the sternum and the other was located 30 cm away in a direction perpendicular to the line determined by the two J-11s. This was midway between the tank's center and the sidewall. The near hydrophone was intended to measure the surface pressure on the subjects' lungs and the further hydrophone was intended to measure the scattered pressure in a way that would integrate over a larger area in the event that lung motion was not uniform. The further hydrophone needed to be located sufficiently far from the tank boundary, which was extremely soft, that a reasonable SNR could be achieved. The third transducer was a PCB™ 352C accelerometer (100 mV/g) that was taped over the subject's fifth intercostal space on the right side and oriented radially. The accelerometer was encapsulated in syntactic foam to make the unit neutrally buoyant. Thus the accelerometer moved in unison with the surrounding fluid, so the tape served only to ensure the location and orientation of the accelerometer rather than to provide mechanical coupling. The fourth transducer was an ultrasonic device that interrogated the motion of the lung surface through the fifth intercostal space on the subject's left. This device employed a pair of 10 MHz transducers: one functioning as a transmitter and the other as a receiver. The received signal was reflected from the lung's surface and phase modulated by the motion of the surface. A similar device has been described by Cox and Rogers.¹⁵ In this implementation

the surface displacement was extracted in real time by demodulation with a phase-lock loop (PLL). This permitted the extraction of phase information that was not possible in the technique employed by Cox and Rogers. The PLL also introduced the dominant noise source into the measurement, which significantly degraded the resolution of the system. Since the experiment reported here, the resolution of this system has been substantially improved by using a direct homodyne demodulation scheme in lieu of the PLL.¹⁶ For each measurement the subject was required to adjust the position of the ultrasonic transducers so that the ultrasonic receive signal was dominated by his lung's reflection. Subjects were aided by a real-time underwater display of the received signal. Some of the subjects proved to be less skilled at this than others and as a result the ultrasonic data are less complete than the data from other measures. Similarly, the failure of two accelerometers under pressure caused the accelerometer datasets to be less complete than the hydrophone data.

IV. DATA ACQUISITION AND PROCESSING

Data were acquired by a direct digitization of the time domain outputs of all four measures during the presentation of each incident signal using a 12-bit digitizer sampled at 8 kHz. Up to ten incident signals were presented to each subject at each depth depending on the available time and the subject's ability to adjust the ultrasonic system to produce a good carrier signal at each depth. Each signal was presented as a series of 16 half-second-long swept-frequency chirps with half-second separation. Although a single chirp of similar duration would have offered a better SNR for measuring a time-invariant response, this mode of presentation offered advantages for the post-processing of the data, including the removal of transients that occurred when the subjects moved or exhausted bubbles. Transients were removed by visual inspection of the data and the remaining signals were averaged. Some noise sources were readily apparent from the raw data. The heartrate of the subjects was apparent in several measures and, in particular, in the ultrasound data. Power line harmonics were also apparent in the data throughout the band of interest (60, 120, and 180 Hz, etc.). These were dealt with by a time-domain subtraction algorithm in which the power line signals were sampled in the quiescent intervals of the raw data. These were then comb filtered, extrapolated across the signal window, and subtracted from the signal before averaging. The averaged signals were converted to the frequency domain by means of a Fast Fourier Transform (FFT) and normalized by incident spectra to form transfer functions. Incident signals were measured at each depth with the subject absent during a compression-decompression cycle that immediately followed the measurement cycle with the subject. This was necessary because J-11 transducers exhibit a depth-dependent resonance at the low end of the band of interest (about 20 Hz when the chamber was at surface pressure).¹¹ At frequencies below this resonance, their output falls sharply, and the two transducers are poorly matched to each other. Incident levels were computed using the manufacturer's calibration constants for the hydrophones. These values were only relevant to the interpretation of the accelerometer and ultrasound data since the same hydrophones

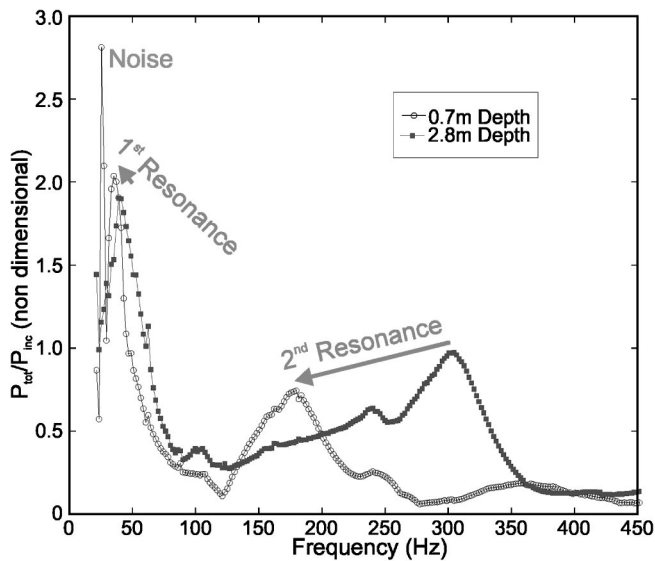


FIG. 4. Magnitude of full bandwidth total-pressure transfer functions at shallow depths from close hydrophone data.

were used for both incident and response measurements. The limited bandwidth of the chirps that were used resulted in a U-shaped noise floor for the measured transfer functions (the actual noise divided by the incident signal spectrum). The J-11s' resonance caused the lower leg of this U to shift upward in frequency with increasing depth. A second problem associated with this resonance was that the poor matching of the J-11's below the resonance frequency caused the incident field to have a significant velocity component that was not accounted for in the lung transfer functions. The amplitude and phase of the pressure associated with the resonant response of the J-11's was normalized out of the lung response when the transfer functions were formed.

Two transfer functions that were measured on a single diver at two different depths are plotted in Fig. 4. These exhibit several features relevant to the further analysis of the data. The low-frequency noise floor can be seen to mimic a resonance (or multiple resonances) in the amplitude of the transfer function. This demonstrates the importance of the inclusion of the measured phase in distinguishing between noise and resonance. There is not a substantial shift in the lower leg of the U-shaped noise floor between the two plotted measurements. The larger low-frequency noise spike associated with the shallower depth (the opposite of the previously mentioned trend) is purely a feature of the noise that was present during that measurement. A higher-frequency resonance is apparent in both transfer functions and it clearly has a stronger depth dependence than the first resonance. This can be partially explained if such resonances are attributed to gas bubbles in the gastrointestinal (GI) tract that shrink in size in addition to increasing their internal pressure and hence their stiffness with depth. An additional effect such as the dissolution of the gas in the bubble or a change in its shape or location must also be postulated to account for the magnitude of the observed changes. It should be noted that a GI bubble need not have achieved equilibrium size for either of these two datasets because gas dissolves over a long time scale and measurements at higher pressures preceded

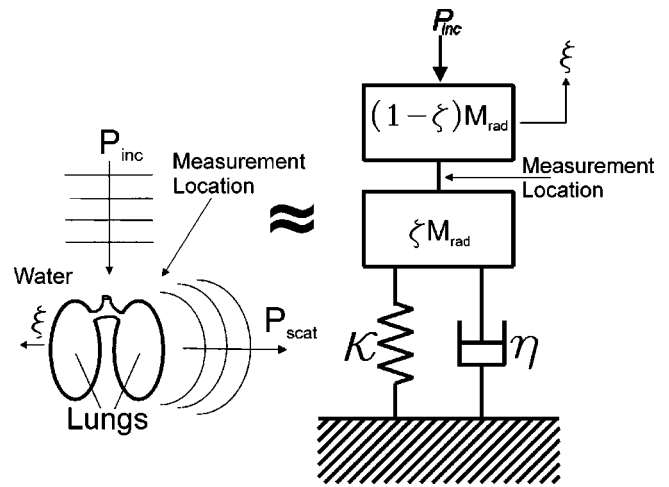


FIG. 5. A one-degree-of-freedom model used for parameter estimation from the measured lung response near the fundamental frequency.

those that are depicted (the order of measurements was from deepest to shallowest equivalent depths). The lung, in contrast, maintains a constant volume because the subject is supplied with air at his ambient pressure through the diving regulator. The second resonance narrows the band in which the lung's response can be evaluated without the specific inclusion of other gas bodies or another plausible physical mechanism to account for this in the model for lung response. It should be noted that the apparently high quality factor ($Q \sim 6$) of fundamental resonances seen in this figure is not real.¹⁷ This is a combined effect of the U-shaped noise floor and the nature of the total-pressure transfer function that does not follow the displacement resonance curve when the pressure is measured anywhere other than directly on the surface of the lung. These effects were considered in the analysis that follows.

V. PARAMETER ESTIMATION

The lung is modeled as a lumped 1-DOF system with the stiffness (κ), damping (η), mass (M), and force (P) defined per unit of the lungs' surface area. The mass term is dominated by the radiation mass of the surrounding fluid, which would be three times the mass of fluid displaced by the lung if the lung were spherical.¹⁸ This mass may be divided into two rigidly attached masses, as is shown in Fig. 5. Here the *excluded mass ratio* (ζ) is defined as the effective entrained mass of moving fluid between the measurement location and the lungs' surface divided by the total radiation mass. Because the attachment is rigid, the response of the system is unaffected by this construction. The pressure transfer function evaluated at the rigid attachment will contain all of the salient features of the interference between the incident and scattered fields close to the surface provided that an appropriate value for ζ is selected to correspond to the measurement location. It is relatively simple to show that, for a spherical bubble, $\zeta = d/(a+d)$, with d and a the distance from the surface to the measurement point and the bubble radius, respectively. This assumes that $k(a+d) \ll 1$, where k is the wave number in the fluid. No *a priori* knowledge of lung size, shape, or measurement-to-surface standoff is re-

quired for this inversion when ζ is used in lieu of a and d . The model acceleration (7) and pressure (8) transfer functions are as follows:

$$\frac{\ddot{\xi}_{\text{surface}}}{p_{\text{incident}}} = \frac{1}{M_{\text{rad}} - i\eta\omega^{-n} - \kappa\omega^{-2}}, \quad (7)$$

$$\frac{p_{\text{total}}}{p_{\text{incident}}} = \frac{\zeta M_{\text{rad}} - i\eta\omega^{-n} - \kappa\omega^{-2}}{M_{\text{rad}} - i\eta\omega^{-n} - \kappa\omega^{-2}}. \quad (8)$$

Here n determines the damping mechanism: $n = -1, 1,$ or 2 correspond to radiation, viscous, or structural damping, respectively. Equation (7) is unremarkable and appears in different forms in textbooks on dynamics and vibrations.¹⁹ Equation (8), however, has interesting features that should be noted. First, it contains a root (for $\eta \rightarrow 0$) that accurately predicts a range-dependent pressure null in the near field of an air bubble. This null explains the false appearance of an excessively high Q in Fig. 4. Second, there are only three free parameters in (8) because a system under study can be scaled by an arbitrary constant with an identical pressure transfer function. Mathematically, this results from dividing any of the lumped parameters (M_{rad} , κ , or η) through the numerator and denominator of (8). The lumped parameters cannot, therefore, be extracted individually from an inversion of the pressure transfer function as they can from the acceleration transfer function. The third inverted parameter (ζ) describes the measurement location and not the lungs. The lumped parameter model was tested by the inversion of synthetic data generated with a free-field bubble model. An assumption of internal lung-like damping was included in the bubble model (i.e., a complex wave number for air that leads to $Q \sim 2.5$). The inversion yielded resonance frequency and Q within 1% of their true values for synthesized pressure measurements made up to one radius from the surface over a frequency range from dc to three times the fundamental resonance of the idealized bubble.

Lung resonance frequencies and Q 's were obtained by fitting (7) and (8) to the experimentally measured transfer functions over a frequency range from 20 to 120 Hz. Parameter fits were performed numerically by minimization of the rms error using the Nelder–Mead simplex search method implemented with the FMINS m file in MATLAB.^{TM20} The noise floor of the data dictated the lower bound for this band and the upper bound was set at a frequency below most of the observed high-frequency resonances that may have involved GI bubbles. Results were discarded for cases where the search algorithm did not converge, a sensor was known to have failed, a resonance was computed outside the band of the data, or the results were clearly aphysical (such as $M_{\text{rad}} < 0$, $\kappa < 0$, $\eta < 0$, $|\zeta| \geq 1$, or $\zeta < 0$). This occurred most often for datasets with poor signal to noise and was often the result of a noise spike appearing to be a resonance peak. These peaks seldom had a phase consistent with resonance, and erroneous signs of one or more model parameter frequently characterized such noise-contaminated results. The largest number of interpretable results was obtained from the close hydrophone, which was the least noisy measure. Here 76% of the data that were acquired provided meaningful parameter sets:

a total of 205 sets for all subjects and depths. The smallest number of results was obtained from the ultrasonic sensor, which was the noisiest. Here only 20% of the data were interpretable for a total of 54 parameter sets. Ultimately the ultrasonic data was of little use because of its poor signal-to-noise ratios. The second hydrophone and the accelerometer provided 190 and 165 parameter sets, respectively. Representative data and curve fits from a single measurement are shown in Fig. 6. Although this measurement is not noise-contaminated to a degree that would make it uninterpretable, the low-frequency nature of the noise along with its relative magnitude on each of the transducers can be seen. The ultrasound data depicted in Fig. 6 is less noise contaminated than it was for most of the measurements. It has been fit to the model in two ways, both considering and neglecting the phase of the measured velocity. The phase-independent predictions are clearly in closer accord with the other data. This may be indicative of unaccounted-for delays in the PLL circuit, but this effect could not later be reproduced and was not diagnosed at the time of the experiment.

Viscous damping offered the best fit to the data for the three scenarios that were considered (i.e., viscous, structural, and radiation damping). The manipulation of n from Eqs. (7) and (8) shows a degradation in both the mean error of the fit and the standard deviation of the resonances measured at each depth if n is varied fractionally from $n = 1$.

The natural frequencies determined by fitting Eqs. (7) and (8) to all of the data are shown as histograms in Fig. 7. The Q 's of these resonances were found from $Q \cong \sqrt{\kappa M_{\text{rad}}}/\eta$, which is taken from Meirovitch.¹⁹ These are depicted in the histograms shown in Fig. 8. Viewed independently, the mass-normalized damping from the hydrophone data and the damping from the accelerometer measurements showed no obvious depth dependence. This would predict that Q is proportional to ω_n ($\omega_n = \sqrt{\kappa/M_{\text{rad}}}$) and M_{rad} is a constant. It is clear from the figures that Q does not consistently obey this relationship. The reason for this is that the accelerometer and hydrophone predictions are in disagreement over the constancy of M_{rad} and η . Accelerometer data indicate that the depth dependence of the resonance is caused by roughly equal contributions from a stiffness that is monotonically related to depth and a radiation mass that is inversely related to depth in the presence of a depth-independent damping term. In contrast, the pressure measurements indicate that the mass-normalized damping is constant and therefore the radiation mass can only vary in direct proportion to the damping. This is one of several interesting features of the data that point to a second degree of freedom for the lung motion with a natural frequency near this frequency range.

Figure 9 shows a comparison between the fundamental resonances predicted by the 1-DOF model fits and a simple model for their depth dependence. Each of the measures has been plotted separately along with error bars that indicate the standard error of the mean for each of the measures. The agglomeration of the data is an unweighted average of all of the data from the two hydrophones and the accelerometer. The ultrasound data were omitted from this average because of a general bias toward higher frequencies that is a combi-

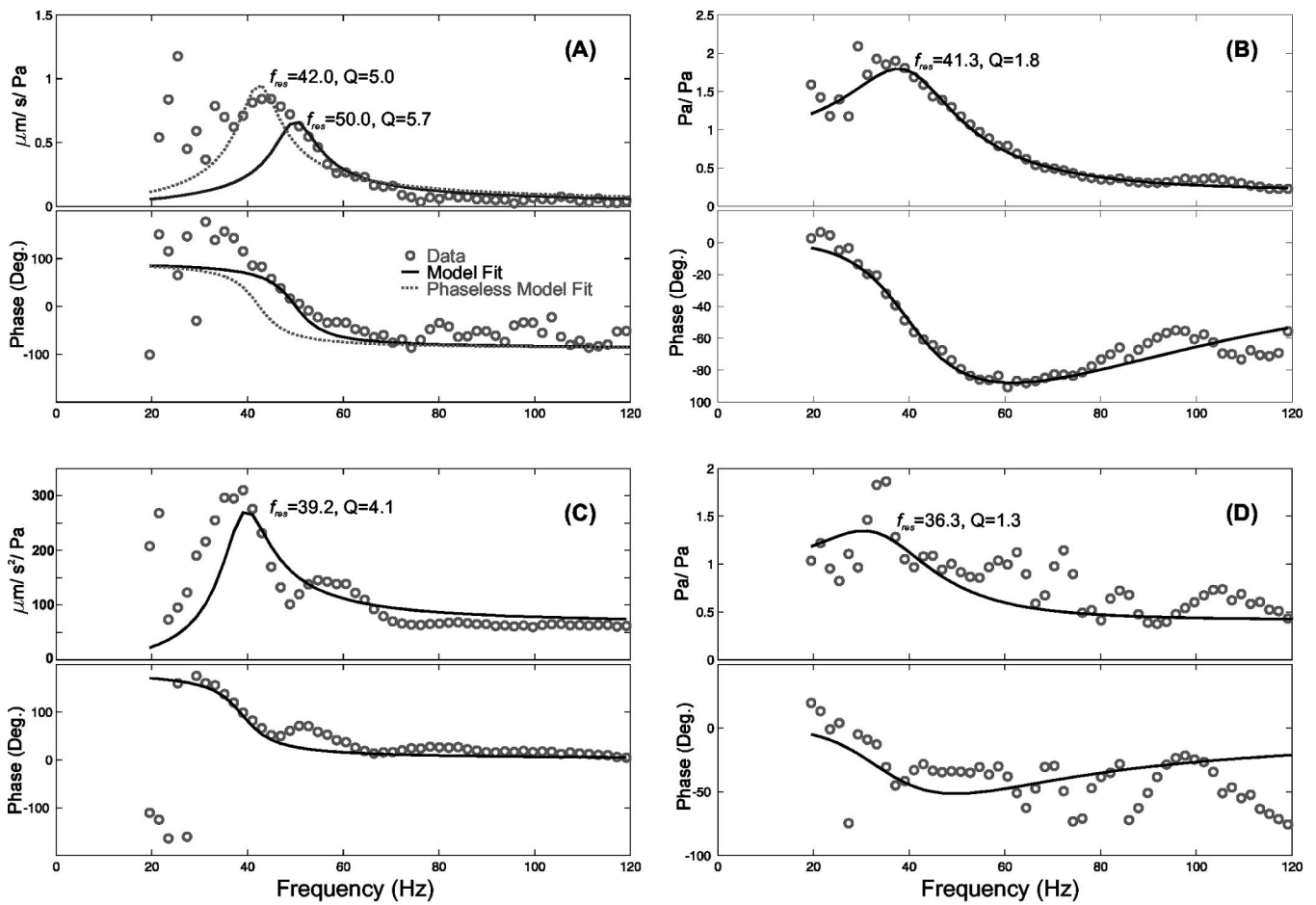


FIG. 6. Full set of transfer functions measured on subject E at 2.8 m equivalent depth: ultrasonic vibrometer (a), close hydrophone (b), accelerometer (c), and far hydrophone (d).

nation of the noise contamination of the data and the phase response of the PLL (as indicated in Fig. 6). The model is based on the assumption of a depth-dependent air spring in parallel with a depth-independent chest-wall stiffness, such that

$$f_{\text{lung}}(z) = f_0 \sqrt{\frac{p(z)}{p(0)}} + \beta, \quad (9)$$

where $f_0 = 31.7$ Hz, $\beta = 0.55$, p is ambient pressure, and z is

depth. f_0 and β were determined from a least-squares fit to the agglomeration of measured resonances. These numbers differ slightly from previously reported results from the same dataset¹⁷ because of the inclusion of additional data and the exclusion of ultrasonic data and aphysical results. In both cases the model fit falls within the error bars for the data to which it is fit. Physically, f_0 represents the natural resonant frequency of the air-spring portion of the lung at atmospheric pressure. β represents the chest wall and body tissue contri-

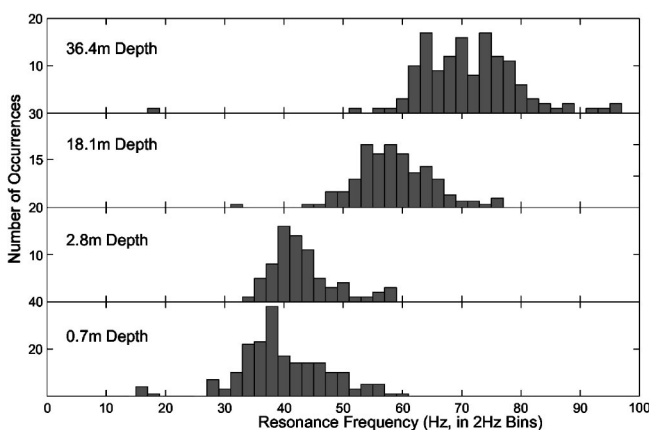


FIG. 7. Histograms of measured resonance frequencies on all subjects for each equivalent depth.

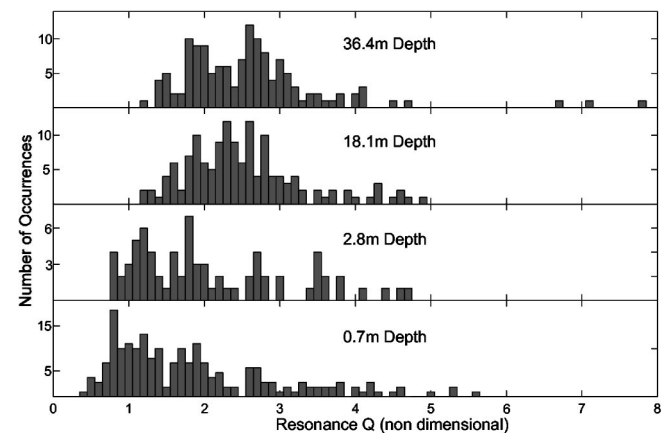


FIG. 8. Histograms of measured resonance quality factors on all subjects for each equivalent depth.

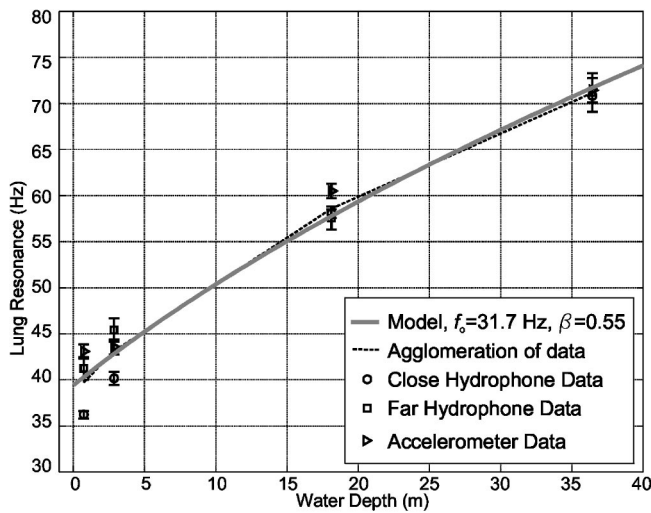


FIG. 9. Data model fit for depth dependence of lung resonance. Error bars indicate standard errors of the mean for each subset of data.

butions to lung stiffness normalized by the air-spring stiffness. The assumption of a depth-independent radiation mass is implicit in this model. This interpretation is reinforced by comparison of f_0 in Eq. (9), as obtained from the data with the natural frequency of a submerged air-filled spherical bubble with the same (4 l) volume. The theoretical value for the resonance frequency of a spherical bubble of radius a is given by¹³

$$f_{\text{bubble}} = \frac{1}{2\pi a} \sqrt{\frac{3\gamma P}{\rho_{\text{water}}}}, \quad (10)$$

where P is the ambient pressure, γ is the ratio of specific heats (1.4 for air), and ρ_{water} is the density of water. For a 4 l volume at one atmosphere ($P = 1 \times 10^5$ Pa) we obtain $f_{\text{bubble}} = 33.1$ Hz that is close to the empirically obtained estimate for f_0 . Equation (9) is also consistent with Andreeva's model for the resonance of a fish swim bladder as an air-filled sphere embedded in an elastic medium:²¹

$$f_{\text{swim bladder}} = \frac{1}{2\pi a} \sqrt{\frac{3\gamma P + 4\mu_f}{\rho_{\text{water}}}}. \quad (11)$$

Here μ_f is the shear modulus of the surrounding medium. Combining Eqs. (9)–(11) reveals that $\mu_f = 3/4\gamma P(0)\beta$. This predicts a shear modulus of 58 kPa for the material surrounding the lung, which is in reasonable accord with reported values for soft body tissue tissues such as blood vessels²² and slightly lower than values reported for soft rubber²³ ($\mu_f \approx 200$ kPa).

The most pronounced deviations from the model fit shown in Fig. 9 are for the measurements made at low ambient pressures. This is somewhat explainable in terms of the salient features of the corresponding datasets. The U-shaped noise floor causes a bias toward higher frequencies in all of the data. This is most pronounced at low ambient pressures because the lung resonance shows stronger depth dependence than the apparent low-frequency noise, which is amplified by the resonant response of the J-11's. Thus, the data from the close hydrophone, which was the most noise-immune measure, result in the lowest mean resonance at low

ambient pressures. Similarly, the data from the ultrasound system (not depicted in this figure) were, in general, very noisy and yielded the highest mean values for resonance frequency. The consistently lower predictions of resonance frequency from the close hydrophone compared to the other two measures depicted in Fig. 9 are interesting. These may indicate additional degrees of freedom. It is clear from Fig. 9 that an explanation that attributes the error to a single type of measure would require an adjustment to the values of f_0 and β in Eq. (9).

VI. ADDITIONAL DEGREES OF FREEDOM

Evidence that additional degrees of freedom play a significant role in the response of the lungs in the frequency range of interest includes systematic disagreements between predictions from different measures with the 1-DOF model regarding the depth dependence of the radiation mass and of the resonance frequency. There is also an obvious high-frequency resonance in much of the data acquired at low ambient pressures. Two different two-degree-of-freedom (2-DOF) models were used to study these effects.

One of these models considers the lungs as two separate systems that are mass-coupled following the simplified models of Shima²⁴ and Zabolotskaya²⁵ for a two-bubble system. Here a complex stiffness was assumed to approximate the Q 's observed in the measured transfer functions. Because the right and left lung differ in size, the new antisymmetric mode has a net volume velocity and can be excited by incident pressure, although it is likely to be more heavily damped than the fundamental mode. This model was used to generate synthetic results that were then inverted using the 1-DOF model that was used to estimate lung parameters from the experimental data. In general, the estimates were good because the two resonance frequencies were close and the low Q 's tended to smear the effects of the two modes together. The inverted parameters reasonably represented the properties of the two bubbles coalesced. There was, however, a systematic difference between the parameters that were inverted from the displacements of the larger and smaller bubbles. This offers one possible explanation for the observed resonance frequency discrepancy between measures. Confirmation of this conjecture would require additional data (such as separate right and left side measures with good SNR) and a more complicated model that accounted for actual lung geometry, nonuniform excitation, and stiffness coupling through the airways. The two-bubble model was also used to simulate the response of GI bubbles. Curves similar to those shown in Fig. 4 were generated by assuming a 3.5 cm diameter bubble in close proximity to a lung-like 20 cm diam (4.2 L) bubble at atmospheric pressure. The response of the larger bubble was negligibly affected by the presence of the smaller bubble below 120 Hz and its parameters could be inverted from the synthesized data in the 20 to 120 Hz range without consideration of the smaller bubble.

A second 2-DOF model that was considered was a mathematical model without an intrinsic physical basis. Here a second resonance (uncoupled from the fundamental resonance) was assumed to exist, and the measured transfer functions were numerically fit to the model. This resolved the

disagreement between hydrophone and accelerometer predictions regarding the radiation mass without significantly altering the fundamental resonances predicted by the 1-DOF model. The modeled second resonance was below the evaluation band and nearly depth independent. The data are best fit by a second resonance at 18 Hz, although the fit can be substantially improved over the 1-DOF model, even by the assumption of a second resonance well below 1 Hz. Its primary contribution to the transfer function is that it adds a constant-phase term in the band of interest (which is in the mass-controlled region of the conjectured mode). Because its effects are most apparent in the accelerometer data, it is likely that this resonance, if it exists, does not have an associated volume velocity and is therefore excited by the non-uniformity of the incident field. The incident pressure has a significant spatial second derivative at all frequencies and a gradient that becomes pronounced below the resonance of the J-11 transducers, where they are poorly matched to each other. Possible modes of a second lung-related resonance at these frequencies include whole-body translation, a Helmholtz resonance, or shell-like modes of the chest wall to which the entrained air does not contribute stiffness. Only the last two of these are consistent with the previously mentioned 18 Hz result. The second low-frequency resonance is a compelling explanation for the anomalous accelerometer response in that it confirms both intuition and hydrophone measurements in supporting a depth-independent radiation mass and it significantly improves the quality of the model fit. There is, however, not sufficient data to verify this conjecture.

VII. CONCLUSIONS

The frequency response of the submerged human lung was measured *in vivo* at several depths over the frequency range of 20 to 500 Hz. Although the measurements were made in a relatively small volume of water, free-field conditions were reasonably simulated with regard to the fundamental lung resonance. Three different techniques were used for the measurement: pressure was measured in the fluid around the subjects at two locations with hydrophones, chest wall acceleration was measured using an accelerometer, and lung surface motion was measured using a noninvasive ultrasonic technique. Noise problems rendered the ultrasonic data useful only as a proof-of-concept. The measurements are consistent with a lung stiffness that is a parallel combination of the chest wall and the entrained air, a dominant viscous damping mechanism contributed by body tissue, and a depth-independent radiation mass that is dominated by the surrounding fluid. Averaged measurements over ten subjects support the conclusion that lungs resonate at around 40 Hz with a mean Q of 1.8 to 2.2 (depending on assumptions regarding the existence of an additional low-frequency resonance) when normally inspired at an ambient pressure of 1.1 atmospheres. The resonance frequency increases as the square root of the ambient pressure and was measured on average at 71 Hz at a pressure equivalent to a 36.4 m submersion in seawater. At this pressure, Q 's rose to about 2.6 on average. A physical model based on the experimental data predicts that the resonance frequency would remain below

the band of interest for the SURTASS LFA system (i.e., <100 Hz) over the range of diving depths down to 86 m.

Although the analysis presumed that the lung could be characterized as a 1-DOF system, there is evidence to suggest that other modes of vibration play a role in the measured data. The more compelling of these are a depth-independent resonance at or below 18 Hz with little or no associated volume velocity that is open to several interpretations and a strongly depth-dependent resonance well above 120 Hz that may be associated with GI bubbles. Because of the small number of sensors used in the experiment, there is not sufficient data available to confirm or characterize the other modes that have been conjectured.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research under Contract No. N00014-97-10-0949. The authors would like to express their thanks to Dr. Robert Gisiner, Dr. Eric Hanson, Dr. Gregg Larson, and to the staff of the NS-MRL hyperbaric chamber facility, who participated in the experiment.

¹ Department of the Navy, "Final Environmental Impact Statement for Surveillance Towed Array Sensor System Low-Frequency Active (SURTASS LFA) Sonar," <http://www.surtass-lfa-eis.com/>, 2001.

² S. Parvin, J. Nedwell, K. Needham, A. Turnpenny, A. Thomas, and S. Searle, "The effects of low frequency sonar transmissions on divers and ichthyofauna: Hyperbaric and open water diver studies and the noise exposure of species of fish," Defense Research Agency, Farnborough, Hampshire, UK, 1994.

³ L. Crum and Y. Mao, "Acoustically enhanced bubble growth at low frequencies and its implications for human diver and marine mammal safety," *J. Acoust. Soc. Am.* **99**, 2898–2907 (1996).

⁴ J. J. Finneran, "Whole-lung resonance in a bottlenose dolphin and white whale," *J. Acoust. Soc. Am.* **114**, 529–535 (2003).

⁵ P. Stein, J. Rudzinsky, M. Binnam, W. Ellison, and J. Johnson, "High frequency marine mammal mitigation active sonar system," *Proceedings of the MTS/IEEE Oceans 2001*, Vol. 3, pp. 1388–1391.

⁶ D. Dalecki, S. Child, and C. Raeman, "Lung damage from exposure to low-frequency underwater sound," *J. Acoust. Soc. Am.* **111**, 2462 (2002).

⁷ C. Clay and H. Medwin, *Acoustical Oceanography: Principles and Applications* (Wiley, New York, 1977), pp. 194–203.

⁸ T. Akamatsu, T. Okumura, N. Novarini, and H. Yan, "Empirical refinements to the recording of fish sounds in small tanks," *J. Acoust. Soc. Am.* **112**, 3073–3082 (2002).

⁹ R. Habib, R. Chalker, B. Suki, and A. Jackson, "Airway geometry and wall mechanical properties estimated from subglottal input impedance in humans," *J. Appl. Physiol.* **77**, 441–451 (1994).

¹⁰ P. Harper, S. S. Kraman, H. Pasterkamp, and G. R. Wodicka, "An acoustic model of the respiratory tract," *IEEE Trans. Biomed. Eng.* **48**, 543–550 (2001).

¹¹ *Underwater Electroacoustic Standard Transducers Catalog*, Transducer Branch, Underwater Sound Reference Detachment, Naval Research Laboratory, May 1982, pp. 129–132.

¹² T. G. Leighton, *The Acoustic Bubble* (Academic, New York, 1994), pp. 136–157.

¹³ M. Minnaert, "On musical air-bubbles and the sound of running water," *Philos. Mag.* **16**, 235–248 (1933).

¹⁴ P. H. Rogers, E. Cudahy, and E. Hanson, *Measurement of Lung Vibration from Low Frequency Sound in Divers*, Experimental Protocol, Naval Bureau of Medicine and Surgery, October 1998.

¹⁵ M. Cox and P. H. Rogers, "Automated noninvasive motion measurement of auditory organs in fish using ultrasound," *J. Vibra., Acoust., Stress, Reliability in Design*, **109**, 55–59 (1987).

¹⁶ J. S. Martin, D. J. Fenneman, F. Codron, P. H. Rogers, W. R. Scott, G. D. Larson, and G. S. McCall, "Ultrasonic displacement sensor for the seismic detection of buried land mines," in *Proc. SPIE: 2002 Annu. Int. Symp.*

- Aerospace/Defense Sensing, Simulation, and Controls*, Orlando, FL, April 2002, Vol. 4742, pp. 606–616.
- ¹⁷J. S. Martin, P. H. Rogers, E. A. Cudahy, and E. L. Hanson, “Fundamental resonance of the water-loaded human lung,” in *Proc. 17th Int. Congress on Acoustics*, Rome, Italy, September 2001.
- ¹⁸L. Kinsler, A. Frey, A. Coppens, and J. Sanders, *Fundamentals of Acoustics* (Wiley, New York, 1982), pp. 228–231 (regarding simple bubble models) and 191–193 (regarding radiation impedance).
- ¹⁹L. Meirovitch, *Elements of Vibration Analysis*, 2nd ed. (McGraw-Hill, New York, 1986), pp. 50–55.
- ²⁰MATLAB Version 5.3.1.29215a (R11.1), The Mathworks Inc., Copyright 1999.
- ²¹I. B. Andreeva, “Scattering of sound by air bladders of fish in deep sound scattering ocean layers,” *Sov. Phys. Acoust.* **10**, 17–20 (1964).
- ²²H. Abe, K. Hayashi, and M. Sato, *Data Book on Mechanical Properties of Living Cells, Tissues, and Organs* (Springer-Verlag, Tokyo, 1996).
- ²³J. M. Gere, *Mechanics of Materials*, 5th ed. (Brooks/Cole, Pacific Grove, CA, 2001).
- ²⁴A. Shima, “The natural frequencies of two spherical bubbles oscillating in water,” *Transactions of the American Soc. of Mech. Engr.*, September 1971, pp. 426–432.
- ²⁵E. A. Zabolotskaya, “Interaction of Gas bubbles in a sound field,” *Sov. Phys. Acoust.* **30**, 365–368 (1984).

Instrumenting free-swimming dolphins echolocating in open water

Stephen W. Martin, Michael Phillips, Eric J. Bauer, and Patrick W. Moore
SPAWAR Systems Center San Diego, 53560 Hull Street, San Diego, California 92152-5001

Dorian S. Houser
BIOMIMETICA, 7951 Shantung Drive, Santee, California 92071

(Received 9 August 2004; revised 8 January 2005; accepted 18 January 2005)

Dolphins within the Navy Marine Mammal Program use echolocation to effectively locate underwater mines. They currently outperform manmade systems at similar tasks, particularly in cluttered environments and on buried targets. In hopes of improving manmade mine-hunting sonar systems, two instrumentation packages were developed to monitor free-swimming dolphin motion and echolocation during open-water target detection tasks. The biosonar measurement tool (BMT) is carried by a dolphin and monitors underwater position and attitude while simultaneously recording echolocation clicks and returning echoes through high-gain binaural receivers. The instrumented mine simulator (IMS) is a modified bottom target that monitors echolocation signals arriving at the target during ensonification. Dolphin subjects were trained to carry the BMT in open-bay bottom-object target searches in which the IMS could serve as a bottom object. The instrumentation provides detailed data that reveal hereto-unavailable information on the search strategies of free-swimming dolphins conducting open-water, bottom-object search tasks with echolocation. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1867913]

PACS numbers: 43.80.Ev, 43.80.Ka [WA]

Pages: 2301–2307

I. INTRODUCTION

Dolphins possess a biological sonar system that is highly effective at submerged object detection and identification, is operational in open water to littoral zones, and is capable of finding buried targets. Due to the ability of the dolphin to out-perform manmade systems at submerged object detection and identification, the United States Navy has placed considerable effort into understanding dolphin biosonar (echolocation).

Past research has investigated the propagation characteristics of echolocation clicks (Au and Snyder, 1980; Au, 1980; Au, 1992), mechanisms of click production and echo reception (Brill, 1991; Bullock and Ridgway, 1972; McCormick *et al.*, 1980; Moore and Pawloski, 1990; Cranford, 2000, Dubrovskii *et al.*, 1991), adaptive production of clicks relative to the echoic task performed (Houser *et al.*, 1999), and performance at object detection and identification (Au and Snyder, 1980; Au and Penner, 1981; Helweg *et al.*, 1996). Dolphin echolocation research is traditionally confined to enclosed pen environments and requires a limitation on dolphin movement and precise placement of the subject to ensure control over the maximum response axis (MRA) of the echolocation beam [see Au (1993) for review of historical studies]. These manipulations are required to manage the complexities of click propagation but impose an artificiality that removes dolphin motion from search strategy analysis, potentially creates a relatively simple echoic task, and minimizes the clutter and reverberation that a dolphin experiences in more naturalistic environments. As such, little quantitative data are available on dolphins conducting free-swimming, open-water acoustic searches.

Prior research conducted in pens or pools is invaluable;

however, it cannot ascertain acoustic interactions with the environment, and the potentially broader variability in response afforded a free-swimming dolphin that conducts an echoic search in a natural, open-water environment. Under such conditions, information obtained on animal movement, click production, echo reception, and reception of clicks at an ensonified target can provide detailed information on dolphin echolocation that can be mined for biosonar search strategies under real-world conditions. Results can be applied to the development of signal-processing algorithms and adaptive search strategies that improve the mine-hunting capability of manmade naval assets (e.g., Au *et al.*, 1995; Gaunaurd *et al.*, 1998; Helweg *et al.*, 1996; Moore *et al.*, 1991; Roitblat *et al.*, 1995).

This paper describes the design of two instruments constructed for free-swimming dolphin echolocation studies in the natural environment. The first instrument, termed the biosonar measurement tool (BMT), permits the movement and echolocation strategy of a target-hunting dolphin to be recorded without hampering the dolphin's motion through the search field. The second instrument, the instrumented mine simulator (IMS), records echolocation clicks at the targeted mine simulator. Collectively, data from the BMT and IMS provide a composite acoustic and visual representation that can be mined for information on dolphin echolocation search strategies.

II. METHODS

A. Biosonar measurement tool (BMT)

The BMT is based on an earlier development effort termed the attitude range monitor (ARM; Sigurdson, 1997). The ARM effort investigated dolphin echolocation during

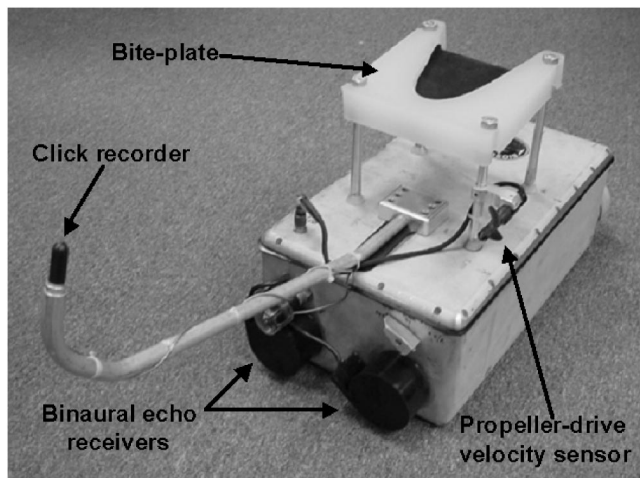


FIG. 1. Photograph of the BMT with the click detector hydrophone, binaural receivers, velocity sensor, and biteplate identified.

open-water target detection trials, but was constrained by requiring that the animal and recording instrumentation be tethered to a research boat. The BMT is mounted underneath a biteplate that the dolphin carries in its mouth (Fig. 1), thus allowing the animal to freely move throughout a search field. This design was favored over alternatives (such as backpack mounting) for simplicity, safety, and comfort of the dolphin. The dolphin is free at any time to open its mouth and release the instrumentation package. The biteplate configuration also provides a consistent mounting relationship between the dolphin and the acoustic and navigational sensors.

The BMT was developed to record a wide variety of sensor data from a free-swimming dolphin engaged in a target detection task. Sensor data include three-dimensional attitudes and dynamics of the dolphin (heading, pitch, roll, acceleration, angular rates, depth, and velocity) and three channels of passive sonar data (one for recording outgoing echolocation clicks, and two for recording echoes on paired biologically inspired high-gain binaural receivers). A low-risk approach to the development of the BMT electronics was employed through the use of Commercial Off-The-Shelf (COTS) electronics.

1. Underwater navigation

Underwater navigational data were measured using a mixture of three separate passive sensors. These included a pressure sensor (Micron Instruments, Simi Valley, CA) for measuring depth to an accuracy of ~ 8 cm and a nonmagnetic, fiberoptic, propeller-driven velocity sensor (Swoffer Instruments, Seattle, WA). Additionally, an all-attitude Crossbow® solid-state, 9-degree of freedom, Attitude Heading Reference Unit (AHRS) was used. The AHRS computed Euler angles of heading, pitch, and roll, and provided three axes each of magnetic data, acceleration data, and angular rate data. The AHRS angular rate sensors required “zeroing” on a stable platform prior to each power on use. AHRS data were used to reconstruct and analyze the dolphin’s underwater track during a search procedure and to investigate more detailed aspects of dolphin motion, such as tail thrusts (e.g., Johnson and Tyack, 2003). Due to inherent bias and stability errors associated with solid-state accelerometers, a propeller-

driven velocity sensor was required to directly measure velocity. Collectively, AHRS data combined with the depth sensor and velocity sensor data allow for estimates of x , y , and z positional data.

2. Acoustic sensors

Three channels were designed into the BMT to record acoustic data—one low-gain channel for sensing echolocation clicks produced by the dolphin, and two high-gain directional channels for sensing echoes from objects in the environment. All channels of acoustic data were continuously and simultaneously sampled to 16-bit resolution at a sample rate of ~ 314 kHz. Sigma-delta analog to digital oversampling converters were utilized to provide a frequency bandwidth of ~ 150 kHz for each channel. All channels were bandpass filtered at 12 kHz on the low end (to reduce the impact of ambient noise in the bay) and 150 kHz on the high end (for antialiasing).

Outgoing echolocation clicks were detected using a small, omnidirectional hydrophone (Reson TC4013) placed 1 m in front of the animal on the presumed maximum response axis (MRA) for outgoing echolocation clicks (Au *et al.*, 1986). This allowed echolocation clicks to be recorded unsaturated to a ceiling of 215 dB re : 1 μ Pa (peak–peak). Fidelity of data collection was maintained for each collected echolocation click by fixing the hydrophone on the presumed MRA, i.e., click variation due to movement of the recording hydrophone was minimal because the position of the hydrophone relative to the dolphin was controlled by the biteplate mounting.

Two directional hydrophones were mounted on the front of the BMT package to receive ambient acoustic data and echoes from objects ensounded by the dolphin’s echolocation. The directional hydrophones were designed to be biomimetic: the hydrophones had beamwidths comparable to behaviorally measured dolphin receive beam patterns (Au and Moore, 1984), and the spacing between biomimetic receivers, 12.5 cm, approximated the spacing of the dolphin’s auditory bullae as determined through computed tomography. The gain of the biomimetic receiver was set such that digitized data saturated at approximately 155 dB re : 1 μ Pa (peak–peak). Due to the wide dynamic range realized with 16-bit converters, and the relatively short functional range of the sonar system (under 100 m), time-varying gain was not utilized. The higher gain and directivity index associated with the directional receiver allowed for the detection of small target strength targets out to ranges in excess of 60 m. The binaural nature of the receiver permitted coherent-based processing, such as beamforming.

3. BMT controller

COTS electronics were utilized for the control computer, which consisted of a conduction-cooled 6U VME computer board with a Motorola PowerPC processor. The VxWorks Real Time Operating System (RTOS) was used for the software architecture. Another COTS 6U VME analog to digital converter board was utilized for acquiring the three channels of acoustic data. Custom electronics were required for

preamplification and signal conditioning of the acoustic channels and the depth and velocity sensors. A battery pack utilizing 10 D-size cells was used to power the unit via a COTS PC104 dc–dc converter. The converter had a switching frequency of 200 KHz, which is above the highest biosonar frequency of interest and well suited to the analog to digital converters utilized. Total power consumption was approximately 35 watts and the device was capable of running for more than an hour on one battery charge. In excess of 20 target search trials were possible on a single battery charge. The batteries were housed in a set of sealed compartments, physically separating them from the BMT controller electronics and facilitating the changing of the batteries in the field. Following a search trial, a wet mate-able underwater connector was utilized to connect the BMT to a computer on the workboat. Data were uploaded to the workboat computer over a standard 100 BaseT Ethernet connection. Time to upload the maximum amount of data collected per trial (242 MB) was approximately 25 s. The BMT was placed in a standby power mode after data upload and prior to the start of the next trial (i.e., in between data collections).

4. Workboat computer

The workboat computer interfaced to a differential global positioning system (DGPS) and implemented a user “mark” system for logging the position of the boat and targets in the search field. The position of the boat was marked when the dolphin carrying the BMT was sent from the side of the boat to search for a target, and was marked again once the dolphin returned to the boat following completion of the target search. Send and return marks allowed integration of the underwater relative navigational data into the world coordinate system (i.e., latitude/longitude). The workboat computer utilized ONTRAK© navigational software, which could be integrated with uploaded navigational data from the BMT to assess the search path of the dolphin.

B. Instrumented mine simulators (IMS)

The instrumented mine simulators (IMS) used in dolphin bottom-object searches were developed approximately 18 months after initial development of the BMT. These instruments made use of a single omnidirectional receiving hydrophone (Reson TC4013A) to collect echolocation signals produced by the dolphin. The gain on the hydrophone preamplifier was set at 40 dB [incoming clicks saturated at ~ 165 dBre:1 uPa (peak–peak)] so that incoming echolocation signals could be recorded when the dolphin was tens of meters away from the IMS. The same sigma-delta analog to digital converter chips used in the BMT were used in the IMS. However, as the IMS is stationary during trials, it contained a relatively simple COTS heading, pitch, roll sensor and no depth or velocity sensors. A COTS PC104 + Pentium-based control computer and data acquisition boards were used to reduce the footprint and expense of the IMS. A SHARC-based PC104+DSP card was used for acquiring single-channel data at a sample rate of 312.5 kHz. A D-cell alkaline battery pack and PC104 dc–dc converter were utilized to power the IMS.

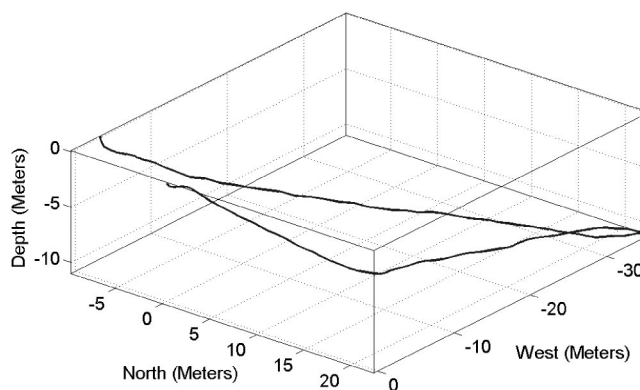


FIG. 2. A three-dimensional representation of a dolphin target search relative to the start-of-search position. The axes units are in meters relative to the starting position.

The IMS included an acoustically commanded wake/sleep subsystem that allowed the IMS to be placed in the field for weeks at a time; in standby mode, an IMS could stay deployed for up to 3 weeks. Each IMS could be “awakened” before conducting dolphin search trials with a 32-kHz tone. The IMS would then continuously collect acoustic data for 18 min before automatically shutting off. IMS packages stored data on a 20-GB hard drive until they were retrieved from the field. Following recovery, data collected with the IMS were transferred to analysis computers utilizing a 100 BaseT Ethernet link.

III. DATA ANALYSIS AND ANALYSIS TOOLS

A. BMT—underwater navigational data

The collective outputs of the BMT navigational sensors can be used to create three-dimensional paths of a dolphin during an echolocation search. Coupled to acoustic data, these reconstructions permit relationships between echolocation production, target ensonification, and the spatial association between target and dolphin to be analyzed. This information is particularly useful during target-present trials in which the dolphin whistles to signify its finding of the target (see Houser *et al.*, this issue). The whistle, which is recorded by the BMT acoustic sensors, permits the acoustic information potentially processed by the dolphin prior to decision making to be segmented from the data stream. This reduces the amount of data that requires postprocessing if one is investigating the target detection and classification aspects of the dolphin’s echolocation.

Figure 2 provides a three-dimensional plot demonstrating the underwater position of a dolphin during a target search as derived from the AHRS, depth, and velocity sensors. Data corresponding to the start of the trial are referenced to the origin of a three-dimensional Cartesian coordinate system. Additional navigation data may also be plotted with position for finer analysis. Figures 3(a) and (b) demonstrate the velocity of a dolphin during a target search (in m/s), as well as the pitch angle of the animal and its acceleration in the Z dimension. Such information is useful in determining the tail thrust sequences during animal motion as well as when the animal demonstrates angular head motion, i.e., when the animal changes the region of echo inspec-

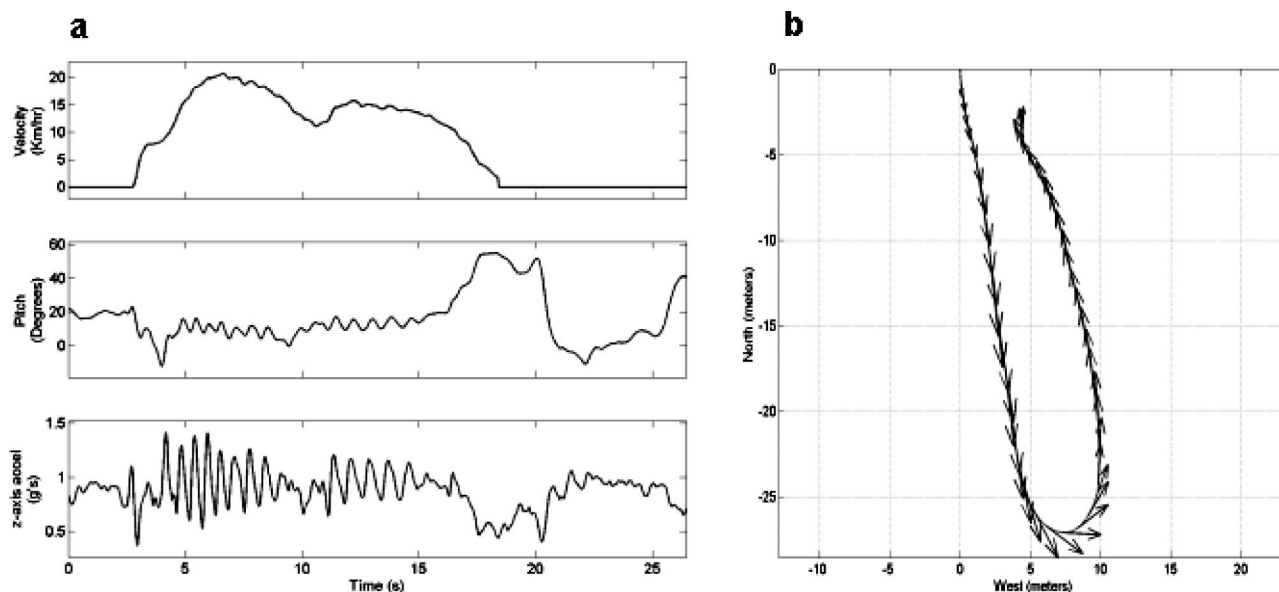


FIG. 3. (a) From top to bottom—velocity in meters/s; middle—pitch angle in degrees; bottom—Z-axis accelerometer in G’s for a dolphin on a correct, target-present trial are shown. A strong effect of tail thrust in the vertical accelerometer (Z axis) can be observed corresponding to the strongest and fastest tail thrusts of the animal in the first 1/3 of the trial. (b) A planar view of travel path in meters with relative velocity demonstrated by vector arrows. The data are from the same trial demonstrated in Fig. 2.

tion without varying the position of the longitudinal axis of its body. Figure 3(b) demonstrates the two-dimensional search of the dolphin and includes velocity information subsampled from the original AHRS data set.

B. BMT—acoustic data

Figure 4 provides an overview of the three acoustic channels of data for the trial presented in Fig. 3 (45-s trial). Figure 5 demonstrates the source level and interclick intervals for all outgoing clicks from a representative short trial.

The trial was a target-present trial and the arrows show where the dolphin whistled to indicate detection of the target. Instantaneous peak source levels are over 205 dB *re*:1 uPa and the interclick intervals (measured as the time between the two pressure peaks of successive clicks) are around 40 ms at the time of the whistle. These graphics demonstrate the utility of the collective data in exploring dolphin echolocation strategies; temporal and spectral characteristics of recorded clicks can be compared with subject attitude and position relative to the target, up to the time that the dolphin whistles to signify target detection.

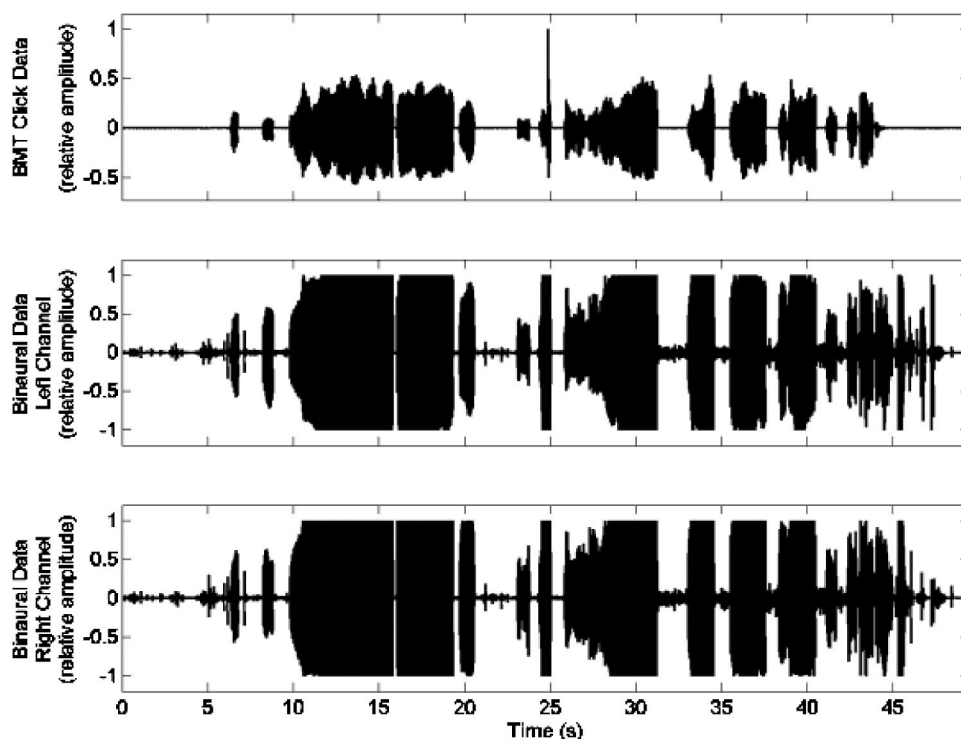


FIG. 4. Acoustic time-series data for a 45-s trial (same trial as Fig. 3). Data from the outgoing click monitor (a) and binaural high-gain, directional receivers (b) are displayed. Note that saturation of the binaural receivers requires use of the time-aligned click channel to segment echolocation clicks from the echo stream.

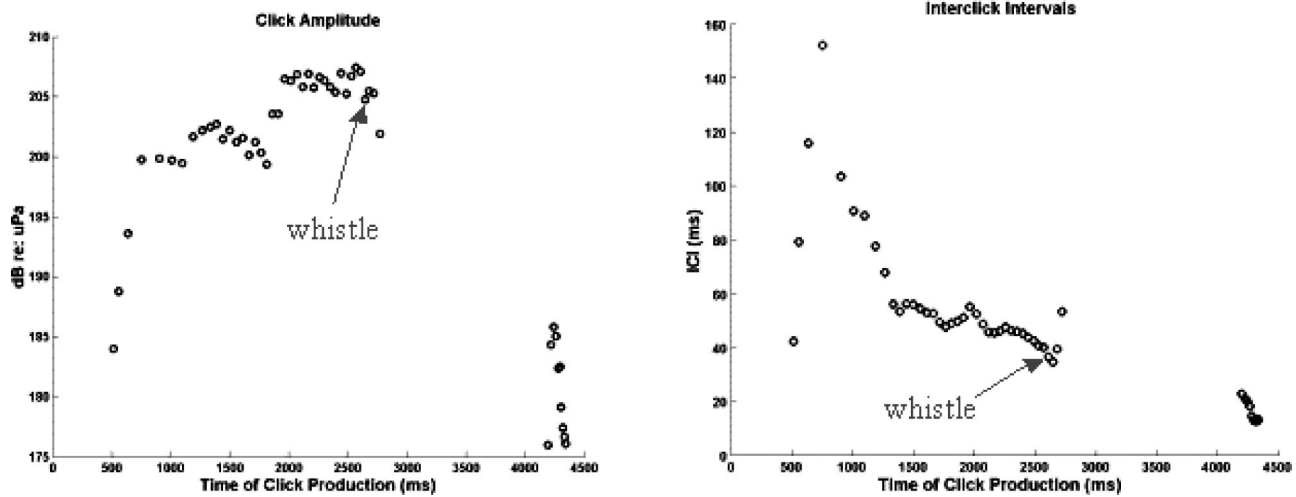


FIG. 5. Peak amplitude in dB re: $1 \mu\text{Pa}$ and interclick interval in ms for the target search depicted in Fig. 3. The arrows indicate the point at which the dolphin whistles, indicating detection of the target.

The acoustic and navigation data can be time aligned by correcting for the start of acquisition for each sensor, thus allowing position and acoustic information to be synchronized. Time alignment is essential for segmentation of the outgoing echolocation clicks from the more sensitive binaural receiver channels (see Fig. 4). One can also gain exceptional situational understanding for each trial by coupling scientific visualization techniques with an acoustic replay of the echolocation data. Echolocation signals can be time stretched into the audible range by slowing the replay by a factor of 8. The three-dimensional dolphin position data, target position data, acoustic data streams, and timing information can then be combined to form a four-dimensional acoustic and visual representation of the dolphin search strategy. Visualization can be augmented with bathymetric features to aid in understanding the dolphin's search strategy in relation to the bathymetry and target placement (for example, see Fig. 6). Various modeling languages can be used for this type of visualization, including the VIRTUAL REALITY MODELING LANGUAGE (VRML) used in Fig. 6. VRML is interactively viewable via a Web browser through freely available plug-ins, such as Cosmo Player[®]. Once acoustic and position data are time aligned, analysis typical to echolocation studies (e.g., variation in amplitude and spectral content of clicks, interclick intervals, etc.) can be applied with added contextual information on the movement of the dolphin relative to the target.

Data collected from the binaural receivers allow backscatter produced by the dolphin's echolocation clicks to be observed. Echoes can be analyzed for temporal and spectral cues that may indicate to the dolphin the presence of a target. Because of the interaural time of arrival differences in the binaural model, it also permits beamforming, or bearing estimates, to determine the relative angle to echo sources. Standard cross-correlation techniques or more elaborate adaptive techniques (e.g., BioSAS[®], Chirp Corporation, La Jolla, CA) may also be applied to investigate echoic returns for clues to target detection. Information from these various analyses can be used to determine if a correlation exists between the frequency-dependent energy distributions of click

and echo returns and the animal's searching mode (i.e., search, acquire, confirm).

C. IMS data

Similar to navigation and acoustic data collected with the BMT, acoustic data collected with the IMS can be synchronized to BMT data after accounting for timing offsets between the instruments. This is accomplished by roughly aligning the time of day based upon internal clocks and then correlating interclick intervals recorded at the IMS with interclick intervals recorded from the BMT. Figure 7 provides a sample of acoustic information arriving at an IMS that has been synchronized with BMT data. The upper panel shows

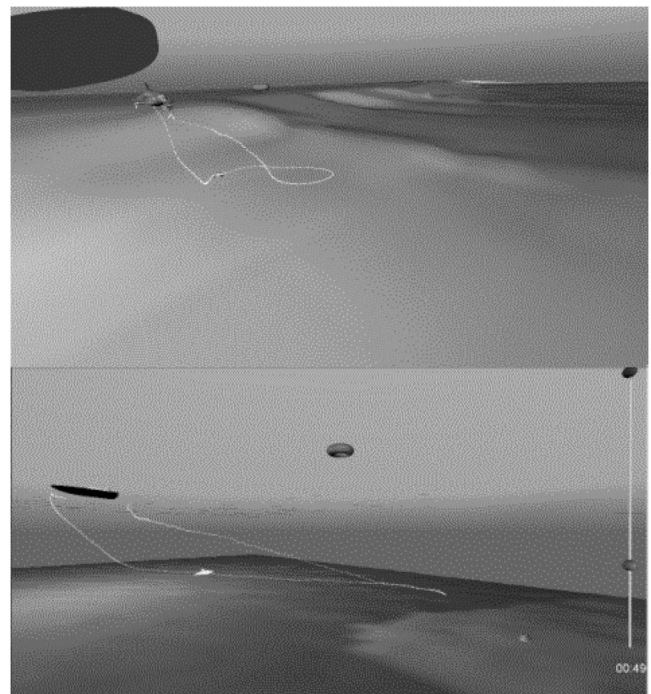


FIG. 6. Interactive four-dimensional representation of the dolphin search depicted in Fig. 3 created using VRML. The viewer used here is CosmoPlayer[®].

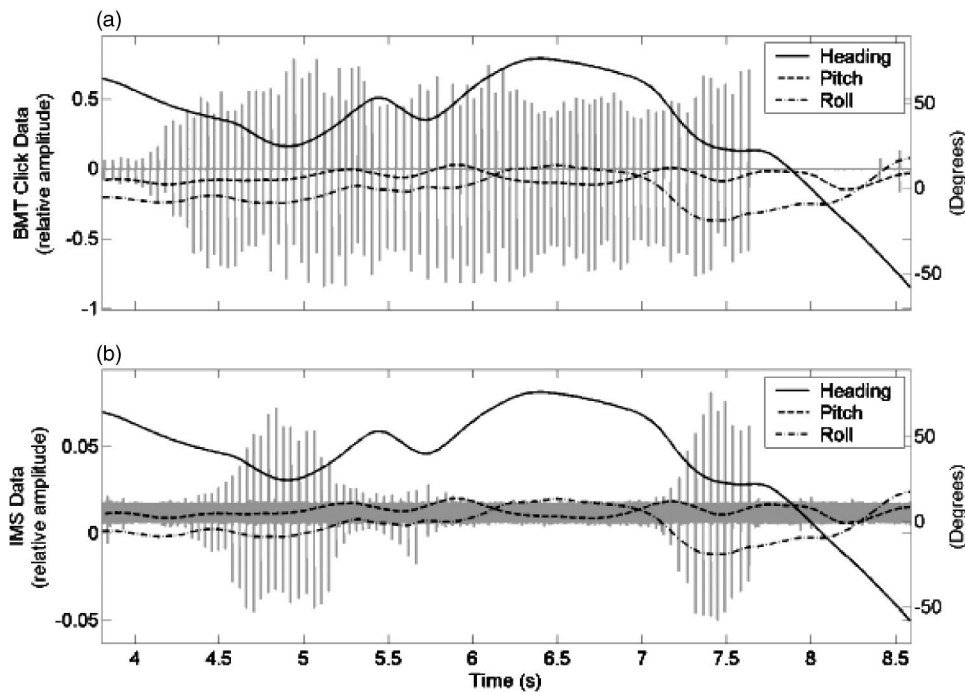


FIG. 7. (a) BMT outgoing click synchronization with (b) IMS incoming clicks, both with respect to BMT heading, pitch, and roll data. The pitch and roll data are centered around 0 deg, while the heading data varies from approximately 140 deg down to 40 deg while the dolphin is searching.

the time series for echolocation click production as recorded by the BMT, as well as the simultaneously collected heading, pitch, and roll. Given the known relative position of the target to the dolphin at the beginning of the trial, bearing estimates from the BMT can be used as an additional step in verifying that the dolphin is inspecting the IMS. The lower part of Fig. 7 shows the IMS incoming acoustic data with the same time-aligned BMT heading, pitch, and roll data that appear in the upper panel. In this instance, the IMS is at a true bearing of approximately 50 deg from the location of the BMT at the time of echolocation click reception. The dolphin acoustic search pattern in terms of head motion can be observed as scanning from a 100 deg heading down to 50 deg. The IMS received signals are maximized when the BMT is pointing at a heading of 50 deg. The dolphin then scans back up towards 100 deg and comes back to 50 deg; again, the IMS data show maximal incoming signal as the dolphin scans back onto the IMS target heading. The acoustic search ends at this time.

Acoustic data from the IMS are useful in determining how many of the dolphin's echolocation clicks ensound the target. This information permits the scope of acoustic information analyzed for cues to target detection to be narrowed and provides insight into the attenuation of frequency-specific energy with click propagation. Coupled to the returning echoes recorded at the BMT, a complete acoustic circuit is observed demonstrating the type of acoustic information available to the dolphin and how that information changes as a function of range and angle of target incidence.

IV. CONCLUSION

The BMT and IMS systems permit the study of search strategies by free-swimming, echolocating dolphins operating in real-world environments. Traditionally, echolocation studies relied upon hoop or biteplate stationing to ensure the on-axis collection of echolocation clicks, necessarily con-

straining the potential behaviors of the animal and the complexity of the tasks. The ability to conduct echolocation studies on free-swimming dolphins increases the richness of behavior that can be explored, and permits a more adequate assessment of dolphin search strategies given the limitations imposed by the noise and clutter inherent to a natural environment. Furthermore, such information can be collected without sacrifice of on-axis fidelity of recorded echolocation clicks.

The instrumentation described herein has been designed to provide quantitative data for free-swimming dolphins conducting mine-hunting tasks. Researchers now are capable of having detailed insight into what dolphins actually do when free swimming and hunting for targets. The data allow researchers to better understand how the dolphin effectively accomplishes object detection and identification tasks through acoustic means. Particular to goals of target detection and identification, it allows signal-processing algorithms developed to improve mine-hunting sonar capabilities to be tested on BMT and IMS data. The performance of the algorithms can then be ranked relative to dolphin performance, which is a benchmark for mine hunting in shallow-water and very shallow-water systems (40–200 and 10–40 ft., respectively). To date, data have been collected from two dolphins engaged in free-swimming, bottom-object searches in conjunction with the deployment of both the BMT and IMS systems. Results from the analysis of the BMT data can be found in Houser *et al.* (this issue); however, the richness of the entire data set is unparalleled and requires continued in-depth analysis.

ACKNOWLEDGMENTS

The authors wish to acknowledge the financial support of the Office of Naval Research (Dr. Robert Gisiner, Dr. John Tague, Dr. Doug Abraham, and Dr. Mardi Hastings) and of the Defense Advanced Research Projects Agency (Dr. Alan

Rudolph). The authors also wish to acknowledge San Diego State University students Tim Herrin, Matt Cross, and Andrea Vidal in the development of the BMT and IMS instrumentation and VRML virtual replay capabilities. Thanks go to Robert Floyd of SPAWAR Systems Center San Diego for his consulting on measurements of dolphin echolocation parameters. For the training of difficult behaviors required in this project, the efforts of the SAIC contractor dolphin training crew are particularly appreciated.

- Au, W. W., and Snyder, K. J. (1980). "Long-range target detection in open waters by an echolocating Atlantic Bottlenose dolphin (*Tursiops truncatus*)," *J. Acoust. Soc. Am.* **68**, 1077–1084.
- Au, W. W. L. (1980). "Echolocation signals of the Atlantic bottlenosed dolphin (*Tursiops truncatus*) in open waters," in *Animal Sonar Systems*, edited by R. G. Busnel and J. F. Fish (Plenum, New York), pp. 251–282.
- Au, W. W. L. (1992). "Application of the reverberation-limited form of the sonar equation to dolphin echolocation," *J. Acoust. Soc. Am.* **92**, 1822–1826.
- Au, W. W. L. (1993). *The Sonar of Dolphins* (Springer, New York).
- Au, W. W. L., Andersen, L. N., Rasmussen, A. R., Roitblat, H. L., and Nachtigall, P. E. (1995). "Neural network modeling of a dolphin's sonar discrimination capabilities," *J. Acoust. Soc. Am.* **98**, 43–50.
- Au, W. W. L., and Moore, P. W. B. (1984). "Receiving beam patterns and directivity indices of the Atlantic bottlenose dolphin, *Tursiops truncatus*," *J. Acoust. Soc. Am.* **75**, 255–259.
- Au, W. W. L., Moore, P. W. B., and Pawloski, D. (1986). "Echolocation transmitting beam of the Atlantic bottlenose dolphin," *J. Acoust. Soc. Am.* **80**, 688–691.
- Au, W. W. L., and Penner, R. H. (1981). "Target detection in noise by echolocating Atlantic bottlenose dolphins," *J. Acoust. Soc. Am.* **70**, 687–693.
- Brill, R. L. (1991). "The effects of attenuating returning echolocation signals at the lower jaw of a dolphin (*Tursiops truncatus*)," *J. Acoust. Soc. Am.* **89**, 2851–2857.
- Bullock, T. H., and Ridgway, S. H. (1972). "Neurophysiological findings relevant to echolocation in marine animals," in *Animal Orientation and Navigation*, edited by S. R. Galler (Scientific and Technical Information Office, NASA, Washington, D. C.), pp 373–395.
- Cranford, T. W. (2000). "In search of impulse sound sources in odontocetes," in *Hearing by Whales and Dolphins*, edited by W. Au, A. N. Popper, and R. R. Fay (Springer, New York), pp. 109–156.
- Dubrovskii, N. A., Zorikov, T. V., Kvizhinadze, O. Sh., and Kuratashvili, M. M. (1991). "Feature description of signals and principles of its organization in the auditory system of the bottlenose dolphin," *Sov. Phys. Acoust.* **37**, 485–487.
- Gaunaurd, G. C., Brill, D., Huang, H., Moore, P. W. B., and Strifors, H. C. (1998). "Signal processing of the echo-signatures returned by submerged shells insonified by dolphin clicks: Active classification," *J. Acoust. Soc. Am.* **103**, 1547–1557.
- Helweg, D. A., Au, W. W. L., Roitblat, H. L., and Nachtigall, P. E. (1996). "Acoustic basis for recognition of aspect-dependent three-dimensional targets by an echolocating bottlenose dolphin," *J. Acoust. Soc. Am.* **99**, 2409–2420.
- Houser, D. S., Helweg, D. A., and Moore, P. W. B. (1999). "Classification of dolphin echolocation clicks by energy and frequency distributions," *J. Acoust. Soc. Am.* **106**, 1579–1585.
- Houser, D. S., Martin, S. W., Bauer, E. J., Phillips, M., Herrin, T., Cross, M., Vidal, A., and Moore, P. W., (2005). "Echolocation characteristics of free-swimming bottlenose dolphins during object detection and identification," *J. Acoust. Soc. Am.* **117**, 2464–2473.
- Johnson, M., and Tyack, P. L. (2003). "A digital acoustic recording tag for measuring the response of wild marine mammals to sound," *IEEE J. Ocean. Eng.* **28**, 3–12.
- McCormick, J. G., Wever, E. G., Ridgway, S. H., and Palin, J. (1980). "Sound reception in the porpoise as it relates to echolocation," in *Animal Sonar Systems*, edited by R. G. Busnel, and J. F. Fish (Plenum, New York).
- Moore, P. W. B., and Pawloski, D. A. (1990). "Investigations on the control of echolocation pulses," in *Sensory Abilities of Cetaceans: Laboratory and Field Evidence*, edited by J. Thomas and R. Kastelein (Plenum, New York), pp. 305–316.
- Moore, P. W. B., Roitblat, H. L., and Nachtigall, P. E. (1991). "Recognizing successive dolphin echoes with an integrator gateway network," *Neural Networks* **4**, 701–709.
- Roitblat, H. L., Au, W. W. L., Nachtigall, P. E., Shizumura, R., and Moons, G. (1995). "Sonar recognition of targets embedded in sediment," *Neural Networks* **8**, 1263–1273.
- Sigurdson, J. E. (1997). "Analyzing the dynamics of dolphin biosonar during search and detection tasks," in *British Institute of Acoustics Proceedings of the Symposium on Underwater Bio-Sonar and Bioacoustics*, Vol. 19.

Echolocation characteristics of free-swimming bottlenose dolphins during object detection and identification

Dorian Houser

BIOMIMETICA, 7951 Shantung Drive, Santee, California 92071

Stephen W. Martin, Eric J. Bauer, Michael Phillips, Tim Herrin, Matt Cross, Andrea Vidal, and Patrick W. Moore

SPAWAR Systems Center San Diego, 53560 Hull Street, San Diego, California 92152-5001

(Received 8 September 2004; revised 8 January 2005; accepted 18 January 2005)

A biosonar measurement tool (BMT) was created to investigate dolphin echolocation search strategies by recording echolocation clicks, returning echoes, and three-dimensional angular motion, velocity, and depth of free-swimming dolphins performing open-water target detections. Trial start and stop times, locations determined from a differential global positioning system (DGPS), and BMT motion and acoustic data were used to produce spatial and acoustic representations of the searches. Two dolphins (LUT, FLP) searched for targets lying on the seafloor of a bay environment while carrying the BMT. LUT searched rapidly (< 10 s), produced few clicks, and varied click-peak frequency (20–120 kHz); FLP searched relatively slowly (tens of seconds) and produced many hundreds of clicks with stereotypical frequency-dependent energy distributions dominating from 30–60 kHz. Dolphins amplified target echo returns by either increasing the click source level or reducing distance to the target but without reducing source level. The distribution of echolocation click-peak frequencies suggested a bias in the dominant frequency components of clicks, possibly due to mechanical constraints of the click generator. Prior training and hearing loss accommodation potentially explain differences in the search strategies of the two dolphins. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1867912]

PACS numbers: 43.80.Ka [WA]

Pages: 2308–2317

I. INTRODUCTION

Interest in dolphin echolocation is spurred by curiosity over the mechanisms involved, its ecological significance, and its applicability to the improvement of manmade sonar systems (Moore, 1997; Roitblat *et al.*, 1995, 1989). Past research on bottlenose dolphins (*Tursiops truncatus*) has yielded insight into echolocation transmit and receive beam patterns (Au, 1980; Au and Moore, 1984; Au *et al.*, 1986), the time, amplitude, and frequency components of echolocation clicks (Au, 1980; Au *et al.*, 1974; Houser *et al.*, 1999), and the hearing ability of this species (Au *et al.*, 2000; Brill *et al.*, 2001; Jacobs, 1972; Johnson, 1968). Investigations on the effectiveness of echolocation in object detection and identification (Au and Penner, 1981; Helweg *et al.*, 1996; Moore, 1997) and echolocation strategies used in object detection and identification (Houser *et al.*, 1999) have also been performed, but such experiments are typically confined to pen enclosures or utilize stationing techniques to control the position of the dolphin. Such approaches control for the complexities of the acoustic environment and constrain the possible interpretation of the results.

Dolphins move in the vertical dimension of three-dimensional space to a great degree, and their primary sensory modality is acoustic, not visual, thus distinguishing them from most terrestrial mammals (excepting bats). These characteristics make conceptualization of echo-image representation within the dolphin difficult and impede our understanding of dolphin echolocation signal processing. Collecting data on free-swimming, echolocating dolphins can

provide information on how dolphins utilize echolocation in more dynamic and difficult real-world tasks, providing information on the relationship between motion, click projection, and echo reception, the amount of echoic information required for decision making, and the magnitude and types of noise (e.g., clutter echoes) that must be resolved for target detection. Such information can provide insight on mathematical principles underlying the signal processing of the echolocating dolphin that may not be obtainable with experiments confined to pen enclosures or that employ stationing techniques to constrain dolphin movement.

Three-dimensional movement during echolocation increases the complexity of signal-processing tasks necessary for object detection and identification. However, three-dimensional movement should also increase the effectiveness of echolocation; returns from successive sonar pings projected toward a point in space progressively increase information regarding that location if coupled with changes in projector position (particularly for aspect-dependent targets; Altes *et al.*, 2003, 1998). The adaptive value of motion during sonar operations is evidenced by the development of such systems as synthetic aperture sonar (SAS). It is reasonable to assume that a free-swimming dolphin, which has evolved the ability to track and capture mobile prey in a visually limiting environment, capitalizes on dynamic position within the water column while echolocating to optimize useful information contained in returning echoes. It follows that the ability to record outgoing echolocation clicks, returning echoes, and the three-dimensional orientation of a free-swimming echolocating dolphin's head can provide informa-

tion valuable to the development of signal-processing algorithms seeking to emulate dolphin performance. Furthermore, such information might provide insight into the capability and function of echolocation in ecologically important activities such as foraging and navigation.

This paper describes data collected from a hardware system specifically designed for recording echolocation in a free-swimming dolphin and that leverages anatomical knowledge of dolphin auditory system structure into its design (see Martin *et al.*, this issue). The data collection device, termed the biosonar measurement tool (BMT), consists of a recording hydrophone on axis with the echolocation beam of the dolphin, two receivers spatially separated according to the distance between the auditory bullae of the dolphin, and a Crossbow® 9-degree-of-freedom motion sensor (Attitude Heading Reference Unit), pressure sensor, and propeller-drive velocity sensor. The BMT allows for the continuous recording of echolocation clicks, returning echoes, and measurement of the movement of a dolphin's head while the dolphin is engaged in a free-swimming target detection task. This paper characterizes the echolocation strategies of two free-swimming bottlenose dolphins engaged in similar object detection tasks.

The purpose of this research is to gain a better quantitative understanding of how dolphins search for targets in highly cluttered, shallow-water environments, which provide a more difficult setting for echolocation detection tasks than traditional pen-constrained research settings. It seeks to determine the variation or consistency in search strategies between different animals given similar tasks in hopes of finding information useful in advancing signal-processing approaches to target detection and classification in shallow, cluttered environments. The ultimate goal of this research is to improve manmade sonar systems by enabling effective target searches in shallow-water and cluttered environments and to remove the Navy's need to use dolphins in the same capacity.

II. METHODS

A. Instrumentation

Details of BMT hardware design are given elsewhere (Martin *et al.*, this issue). A more general description of design and functionality are given here. The BMT consists of a data acquisition system compartmentalized within a pressurized housing that is backfilled with argon gas. The housing is suspended from a neoprene-covered biteplate that a dolphin uses to carry the unit. The BMT is activated just prior to sending a dolphin on a submerged, target detection task (see subsection B, behavioral task) and it continuously collects data on echolocation production, echo reception, and motion throughout the detection task. Data are collected from each acoustic channel, one "click" channel and two "receive" channels, at ~314-kHz sample rate with 16-bit resolution. Sigma-delta analog to digital oversampling converters are utilized that provide a bandwidth of approximately 150 kHz for each channel. All channels are bandpass filtered at 12 kHz on the low end and 150 kHz on the high end for anti-aliasing purposes. Data collection is manually terminated

when the dolphin returns from the task and reports that a target is either present or absent. Data are then uploaded to a computer through a 100-baseT Ethernet connection. Through interface with a boat-mounted differential global positioning system (DGPS), this computer also records the location of targets and buoys (see subsection B) and the location of the workboat, from which the dolphin begins and ends its search, at the start and termination of a trial. DGPS locations are used to create a map of the search field through ONTRAK® navigational software. The map is used to obtain a larger situational understanding of dolphin movement relative to the location of trials and targets in a session.

B. Behavioral task

Data were collected from two male dolphins that were trained for open-water target searches, LUT (24 years old) and FLP (25 years old). Trials were run between July 2001 and January 2003 and were performed at two open-water sites located within San Diego Bay, San Diego, CA. The test field for LUT was an approximately 132×99-m rectangle which covered a variable range of depths from a maximum of 23.8 m near the bayward boundary to the shallows of the shoreline. The test field for FLP was an approximately 231 × 112-m rectangle which covered a graded range of depths from 14 m at the bayward boundary to the shoreline. The depth grade at this site encompassed a steep upward slope on the shoreward approach from the center of the bay.

Between 4 and 6 search stations were created at each site prior to a session, with an equal number of target present and target absent stations. Each station consisted of a floating buoy alone (target absent) or a floating buoy and a target (target present). Floating buoys were tethered to the seafloor via a weighted clump. The target consisted of an aluminum ball (10-in. diameter, 0.19-in. wall thickness) attached to a short aluminum pipe and mounted on a aluminum base plate (30-in. square, $\frac{1}{2}$ in.). At target present stations, the target was placed 2–30 m away from the base of the weighted clump at arbitrary bearings. Search stations were placed so that the buoys and targets of different search stations were a minimum of 50 m apart. Search stations were rearranged between sessions to prevent memorization of target locations by the dolphins. Varying the search station positions also permitted a variable number of water depths and bottom topographies to be utilized during the trials. Furthermore, natural influx and efflux of sand, mud, vegetation, and false targets (i.e., trash) provided for an environment that varied dynamically with time.

Dolphins were trained to swim a path around the buoy at each station and to search the adjacent areas for the target. If the dolphin detected the target, it whistled, returned to the boat, and reported to a response paddle corresponding to "target present." If the dolphin did not detect the target, it finished the search path around the buoy, returned to the boat, and pressed a response paddle corresponding to "target absent." Dolphins were not required to swim around the buoy if they detected the target prior to reaching the buoy line. Conversely, a correct response on a target-absent trial required that the dolphin swim around the buoy, thus ensur-

ing that he moved through the entire search space within which a target could be placed.

All trials were run from a Boston whaler. Prior to a session, the BMT navigation sensors (angular rate gyros) were zeroed on a stationary cement piling and the BMT was put into standby mode until the start of the trial. Dolphins either swam next to the boat on the way to the test site or were trained to beach themselves on a foam-padded mat within the Boston whaler and were then transported to the test site. Two trainers were utilized for dolphin control and monitoring throughout the session. A third and fourth person were responsible for boat driving and downloading data from the BMT. Prior to the beginning of a search trial, each dolphin was required to station at a trainer's hand on the aft-port side of the boat. The second trainer placed the BMT just below the water surface on the port side of the boat with the biteplate oriented toward the dolphin. The dolphin was then cued to accept the BMT via the biteplate. Once accepted, the BMT was switched to data collection mode and the dolphin was cued to begin the trial.

The Boston whaler was positioned with the port side toward the buoy marker of a search station. Distance between the Boston whaler and buoy varied from 25–60 m. The distance was varied from trial to trial to provide a variable range of areas over which the dolphin was required to search. The ordering of target-absent or target-present trials was determined by consulting a Gellerman series. The series was balanced to provide an equal probability of absent versus present trials within a session but with no more than three trials of the same type in succession.

C. Analysis

Trials were selected for analysis based upon the following criteria: (1) unambiguous animal response; (2) no failure of system hardware during data collection or data upload; and (3) absence of system or environmental noise prohibiting the detection of echolocation clicks. Trials were segregated according to the dolphin's response to the trial. "Correct" trials were those for which the dolphin responded correctly for a target present or target absent, respectively. "Missed" trials were target-present trials that the dolphin incorrectly reported as target absent. "False-alarm" trials were target-absent trials incorrectly reported as having a target present.

Echolocation clicks were detected by using a 1000-point sliding window instantaneous peak-pressure detector with a 175-dB *re*:1- μ Pa threshold. Once detected, echolocation clicks were recorded as a 256-point waveform, with a 32-point buffer prior to the pressure peak. This threshold level occasionally captured other impulsive sounds (e.g., snapping shrimp) or triggered due to other environmental or system noise. Therefore, echolocation data were manually displayed in sequence so a human observer could verify that captured signals were echolocation clicks. Other captured waveforms were discarded from the data set.

Trial start time was defined as the time that the first echolocation click was emitted. Trial end time for target-present trials was defined as the time at which the dolphin whistled, indicating target detection and identification. Trial end time for target-absent trials was determined as the time

that the last echolocation click emitted by the dolphin was recorded by the BMT. Similar criteria were applied to trials in which the animal missed a target or a false alarm occurred; target-absent and target-present definitions for total trial time were applied to each, respectively.

Interclick interval was determined for successive pairs of clicks by measuring the time interval between the instantaneous peak pressures of each click. Interclick intervals greater than 500 ms were eliminated from the interclick interval data set and arbitrarily defined as a termination point of a click train. Each click was transformed to the spectral domain by applying a 256-point FFT, without windowing, thus producing a frequency spectra representation with a \sim 1.2-kHz bin resolution. Peak frequency, instantaneous peak pressure, and 3-dB bandwidth were calculated for each echolocation click.

Echolocation clicks were categorized according to the classification scheme of Houser *et al.* (1999), which is an extension of descriptions presented by Au *et al.* (1995). Briefly, click amplitude spectra were reduced to a number of features and categorized so that categorical analysis of echolocation clicks between trial types could be performed. Rules for click categorization were modified slightly so as to capture the more salient features of the clicks. Previous categorization of dolphin echolocation clicks resulted in a number of clicks being classified as "multimodal" because the classification scheme failed on parameters that were of secondary importance to those that more effectively described the amplitude spectrum (Houser *et al.*, 1999). Specifically, the following modifications to Houser *et al.* (1999) were made:

- (1) Type B and D clicks were eliminated from the classification scheme. These clicks consist of a low-frequency or high-frequency primary peak (within the -3 -dB band) with a secondary peak in the -10 -dB band. These categories were found to seldom occur in previous work (Houser *et al.*, 1999). The remaining categories used in this study consisted of wideband (W), low-frequency unimodal (LFU), high-frequency unimodal (HFU), bimodal (BI) and multimodal (MULTI) clicks.
- (2) Contiguous regions of the amplitude spectrum within the -3 -dB band were not counted as regions unless the bandwidth of the contiguous region exceeded 10 kHz. (For primary peaks to be calculated for a region, the bandwidth of the contiguously bounded region had to exceed 10 kHz.)
- (3) Two adjacent contiguous regions were considered one region if the frequency separation between the adjacent upper and lower frequency limits was less than 5 kHz.
- (4) The frequency considered for categorization was \sim 23–120 kHz.

Target-absent and target-present trials in which correct responses were given by the dolphin were analyzed to determine whether animals demonstrated systematic approaches to click production while performing target searches. Correct target-present trials were further separated into "search" and "acquisition" segments to assess whether click production varied as the dolphin moved from a search mode to an ac-

quisition mode. Data from the receive channels and click channel of the BMT were used to determine times that bound the two segments. The “acquisition” segment was defined as the time that the echo of the target was first detected to the time that the dolphin whistle was detected in the receive channels. The search segment was defined as the time from the production of the first echolocation click to the time that the acquisition phase began. Returning echoes from the insonified target were detected by application of a whitened and match-filtered (WMF) detector. The WMF detector was implemented by convolving the received signal with the time-reversed transmit replica. The output was half-rectified and filtered via a Hanning window. A signal level was calculated on a 32-sample frame, slipped 50%, as 10 log of the variance of the filtered output. This result was then compared to an adaptive noise level estimator and, when the signal level to adaptive noise level ratio exceeded a threshold value of 8 dB, detection was declared. A human observer with experience in identifying echoes from the specific target employed visually verified detections as originating at the target. Even with a trained observer, verification of the target echo was difficult on many trials, so only those trials for which the target echo could be verified were kept for segment analysis.

The complete number of raw detections was recorded for each animal on each trial to assess the amount of clutter that dolphins encountered during searches. Raw detections were clustered by grouping detections in the receive channel within a time frame that was equivalent to the amount of time required for sound to travel 1 m through ocean water, i.e., range clustered to 1 m, which corresponds to the maximum length of the target of interest considered for this task.

Animal movement was plotted in three dimensions by utilizing the yaw, pitch, and roll information from the Crossbow® 9-degree-of-freedom (9-DOF) motion sensor. Depth information was obtained from a pressure sensor and velocity information was obtained from a propeller-driven velocity sensor (Martin *et al.*, this issue). The pressure sensor was used for actual depth determination, and comparisons were made with the depth profiles calculated using the integrated output of the angular rate and acceleration sensors. Furthermore, it permitted the determination of continued descent by the dolphin during periods where the dolphin was descending but looking upward in the water column. (The 9-DOF sensor is essentially fixed with respect to the animal’s head as a result of biteplate mounting. This allows accurate measurement of where echolocation clicks are directed and the reception of echoic information similar to what the dolphin receives. However, the dolphin often moves in a direction which does not correspond to the direction the head is pointed; e.g., the dolphin can look left or right while traveling in a straight line.)

Three-dimensional track information was combined with known depth values for the test site and known DGPS locations of the boat, buoys, and targets on each trial. Collectively, these data were used to create a virtual representation of the dolphin search using the VIRTUAL REALITY MODELING LANGUAGE (VRML; Martin *et al.*, this issue). The VRML representations were used for three purposes: (1) to allow the

researcher to review the dolphin trial with unparalleled situational understanding (i.e., by allowing echolocation to be listened to while simultaneously observing dolphin motion in three dimensions); (2) to quickly identify interesting sections of a trial for more detailed analysis; and (3) to validate spatiotemporal matching of acoustic data (echolocation clicks and target echoes) with positional data (DGPS and underwater positions of the dolphin relative to the target). VRML representations were particularly useful on target-present trials in which the dolphin detected the target since they allowed the position of the dolphin to be compared to the position of the target at the time the dolphin whistled.

D. Statistical analyses

Several statistical analyses were performed to investigate whether the subjects demonstrated consistency in search strategy between target-present and target-absent trials and between search and acquisition segments of target-present trials. Only trials in which the animals correctly responded to either target presence or absence were used for statistical analysis. Significant differences between the peak sound-pressure levels of emitted clicks, interclick intervals, and the number of clicks emitted during different segments of target-present trials and between target-present and target-absent trials were assessed via t-test. Where data proved heteroscedastic, a t-test for unequal variances was used. Significance tests were run with $p=0.05$. (All statistical values reported for the remainder of the paper are means \pm 1 standard deviation.)

The occurrence of clicks was summed across target-absent trials according to the click taxonomy, i.e., the number of clicks belonging to each taxonomic category was determined for each target-absent trial and then summed across all target-absent trials. This process was repeated for target-present trials and then repeated once more using only search or acquisition segments of target-present trials. Summed occurrences were used to test whether certain categories of clicks were produced more often on target-absent trials than target-present trials, or on search segments than acquisition segments. Since the duration of trials was not a controlled factor, and more clicks of any given type could be produced on long trials or segments relative to short ones, the proportional production of each click by category was calculated to normalize the data sets to time. Proportions were arc-sine transformed according to a modification of the Freeman and Tukey (1950) transformation (Zar, 1984). This procedure corrected for the bimodal distribution of proportional data and is the recommended procedure for proportional data that approach either 0 or 1. Two-way analysis of variance (ANOVA) was applied to the data sets utilizing both segment and click category as factors, and trial type (target absent or present) and click category as factors. Homoscedasticity was assessed through application of the Bartlett–Forsythe test (Brown and Forsythe, 1974). When data were not homoscedastic, the Welch ANOVA was applied. Multiple *post hoc* comparisons were performed with either Tukey’s “honestly significantly different” test (HSD; when data were homoscedastic) or the Kruskal–Wallis rank sum test (when data were

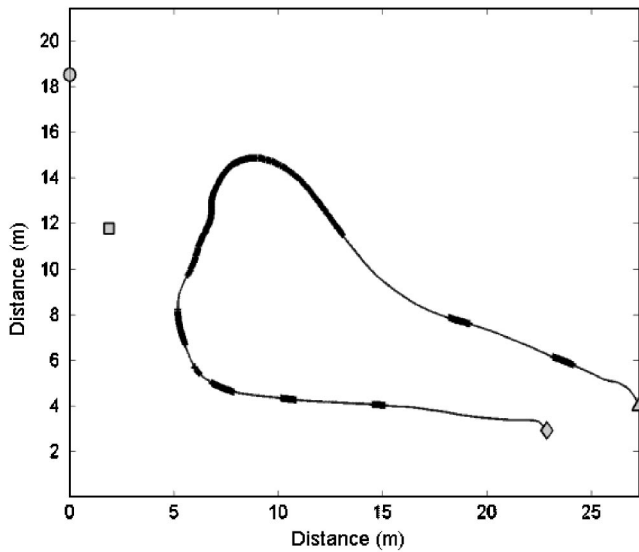


FIG. 1. Demonstration of the search strategy used by FLP relative to an arbitrary starting position. The circle corresponds to the buoy, the square corresponds to the bottom target, and the triangle and diamond correspond to the start and end of the trial, respectively. Individual echolocation clicks are denoted with crosses along the track line. Note that during bouts of echolocation, the density of clicks produced by FLP produces a thickening of the track line.

heteroscedastic) to specify differences in the proportional use of click types. All statistical tests were run with $p = 0.05$.

Categorized clicks were used to construct transition matrices to determine whether bouts of clicks were consistently produced in sequence as part of a consistent search strategy. Click bouts were defined as one or more clicks in sequence that belonged to the same click category. Any categorical change between sequential clicks defined the end of a bout and the beginning of a new bout. (It should be noted that this process potentially overestimates the contribution of certain click types that may have occurred as single transitional clicks between differing click bouts.) The proportion of time that one bout transitioned to another was then tabulated and presented as a confusion matrix.

III. RESULTS

Data collection occurred between July 2001 and January 2003. Collections with FLP were interrupted due to animal health issues, resulting in a much smaller data collection effort with this animal. A total of 431 trials was completed with LUT, while 26 trials were performed with FLP. Application of the criteria used to select trials for further analysis resulted in the following number of trials analyzed for FLP: 7 correct target-present trials, 8 correct target-absent trials, and 6 missed trials. FLP did not report any false positives

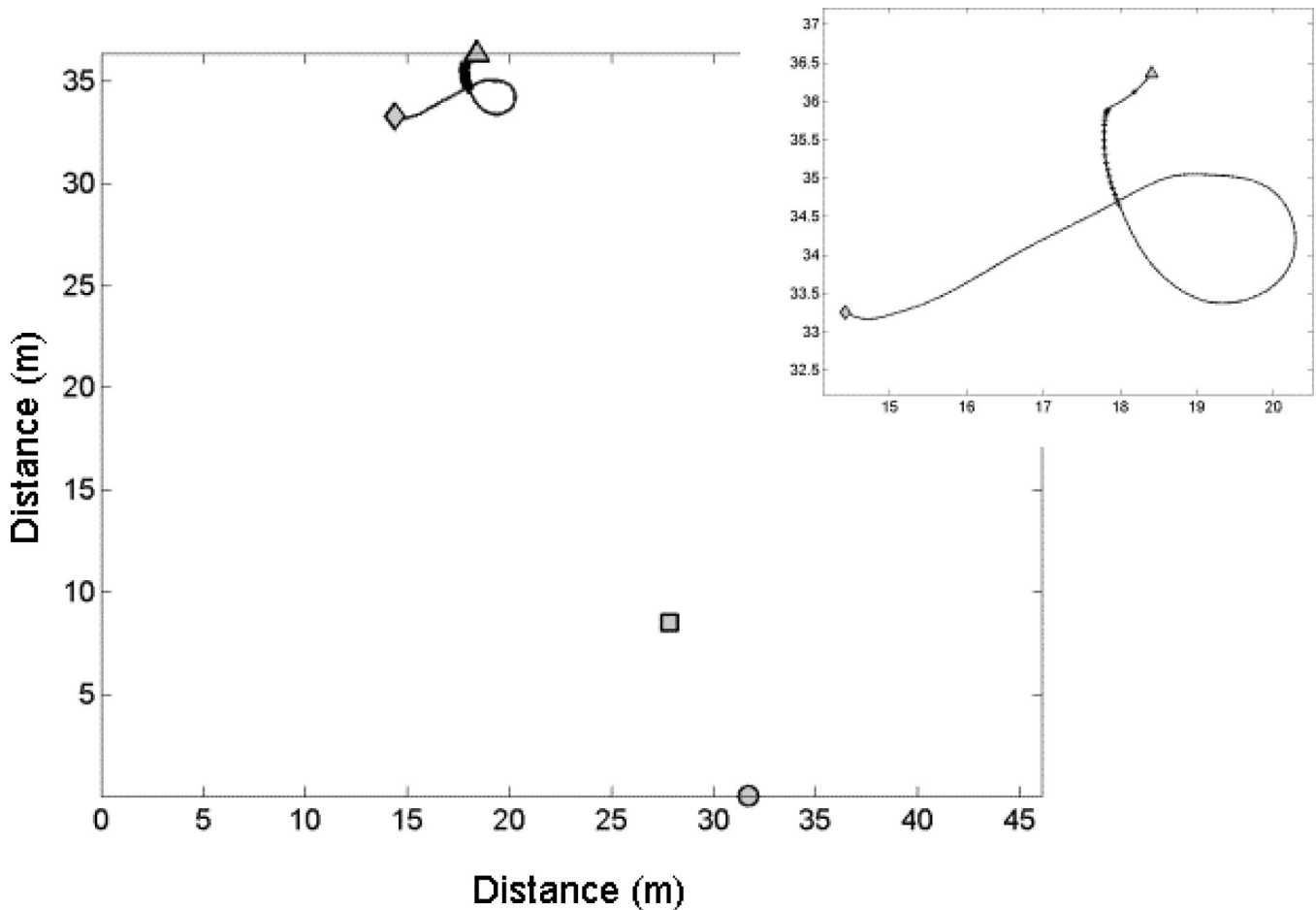


FIG. 2. Demonstration of the search strategy used by LUT relative to an arbitrary starting position. Symbols correspond to those identified in Fig. 1. The panel in the upper right of the figure is a magnification of the search trial, permitting a better view of when echolocation clicks were produced during the trial. Note the much smaller distance traveled relative to FLP.

TABLE I. Comparison of interclick intervals (ICI) between individual dolphins for correct target absent and target present trials and for search and acquisition phases of correct target present trials. Values are means \pm 1 standard deviation and sample size (n).

	FLP		LUT	
	ICI (ms)	n	ICI (ms)	n
Target absent	35.9 \pm 21.9	4437	58.3 \pm 55.1	9412
Target present	38.9 \pm 19.0	2151	61.1 \pm 26.5	2832
Search	38.7 \pm 14.5	1383	66.1 \pm 29.5	652
Acquire	39.3 \pm 25.0	768	56.2 \pm 14.2	407

prior to exclusion from the study. Application of the selection criteria resulted in the following number of trials analyzed for LUT: 96 correct target-present trials, 92 correct target-absent trials, 31 missed trials, and 53 false positive trials. The inability to identify and track target echoes within the receive channels further reduced the number of target positive trials used for categorical comparisons of clicks between search and acquisition segments to 39 for LUT and 7 for FLP.

To characterize the success of the search strategy employed by each dolphin, the probability of detection (P_D) and probability of false alarm (P_{FA}) were calculated for each. P_D was calculated as the ratio of the number of correct responses to target-present trials to the total number of target-present trials. P_{FA} was calculated as the ratio of the number of false alarms to the total number of target-absent trials. FLP had a $P_D=0.60$ but never false alarmed ($P_{FA}=0.00$); in contrast, LUT had a $P_D=0.78$ and $P_{FA}=0.36$.

Differences between animals were noted with respect to the amount of time spent searching during a trial and the movement of the dolphin through the search field. Mean trial time was 24.6 \pm 7.3 s for FLP and 6.5 \pm 3.0 s for LUT on target-present trials. Mean trial time was 38.3 \pm 5.4 s for FLP and 17.6 \pm 5.2 s for LUT on target-absent trials. FLP demonstrated a stereotypical search pattern that consisted of an immediate descent to a depth just above the benthos followed by a thorough and methodical search of the region in the path surrounding the buoy (Fig. 1). In contrast, LUT minimized the distance traveled from the workboat (Fig. 2) and only circled the buoy when a target was not detected. Distance moved away from the workboat on target present trials in which the target was detected was typically 2–3 m for LUT, but was on the order of tens of meters for FLP.

The number of clicks produced per trial, the instantaneous peak pressure of clicks, and the interclick interval significantly varied between target-absent and present trials conducted by both FLP and LUT. FLP produced an average of 312.6 \pm 102.9 clicks per trial on correct target-present trials and 563.8 \pm 83.7 clicks per trial on correct target-absent trials. Click production by LUT was economical by comparison; LUT produced an average of 30.6 \pm 14.7 clicks per trial on correct target-present trials and 108.9 \pm 51.3 clicks per trial on correct target-absent trials. Slight but significantly longer duration interclick intervals were observed during target present trials both for FLP [$F(1,4859)=32.1$, $p<0.001$] and for LUT [$F(1,9937)=13.7$, $p<0.01$; Table I]. Similarly, the peak pressure of clicks emitted on target-

TABLE II. Comparison of click-peak pressure level (dB *re*: 1 uPa) between dolphins for correct target-present and target-absent trials and for search and acquisition phases of correct target-present trials. Values are means \pm 1 standard deviation and sample size (n).

	FLP		LUT	
	dB <i>re</i> : 1 uPa	n	dB <i>re</i> : 1 uPa	n
Target absent	195.2 \pm 8.4	4510	197.5 \pm 6.7	10017
Target present	197.5 \pm 6.8	2188	198.9 \pm 7.3	2942
Search	197.1 \pm 7.4	1412	196.1 \pm 7.7	692
Acquire	198.1 \pm 5.4	776	202.5 \pm 3.7	407

present trials was slightly but significantly greater than the peak pressure of clicks emitted on target absent trials [FLP: $F(1,5208)=138.0$, $p<0.001$; LUT; $t=-9.5$, $df=12957$, $p<0.001$; Table II].

Click production between search and acquisition phases demonstrated less variation between phases for FLP than was observed for LUT. Significantly fewer clicks were emitted by LUT ($t=-2.191$, $df=76$, $p=0.03$) during the acquisition segment (12.3 \pm 6.6 clicks, $n=39$ trials) than during the search segment (15.9 \pm 7.9 clicks, $n=39$ trials) of target-present trials. No significant differences were noted between the number of clicks emitted during search and acquisition segments conducted by FLP (search: 148.9 \pm 55.1 clicks, $n=7$ trials; acquisition: 163.7 \pm 119.9, $n=7$ trials). Interclick intervals ($t=-6.3$, $df=1057$, $p<0.001$) were significantly shorter during acquisition segments than search segments for LUT, but no such relationship was observed for FLP (Table I). The peak pressure of emitted clicks was significantly greater during acquisition segments both for FLP [$F(1,2028)=13.1$, $p<0.01$] and LUT [$F(1,1063)=344.4$, $p<0.0001$; Table II].

ANOVA performed on proportions of taxonomic click types indicated no effect from search or acquisition segment or from the overall trial type, but did indicate a predominant use of wideband clicks throughout the searches conducted by both animals. Results of the Kruskal–Wallis rank sum test are presented in conjunction with the mean and standard deviations of the arc-sine transforms both for FLP (Fig. 3; Kruskal–Wallis SE=84.4, $q_{(0.05,\infty,5)}=3.858$) and for LUT

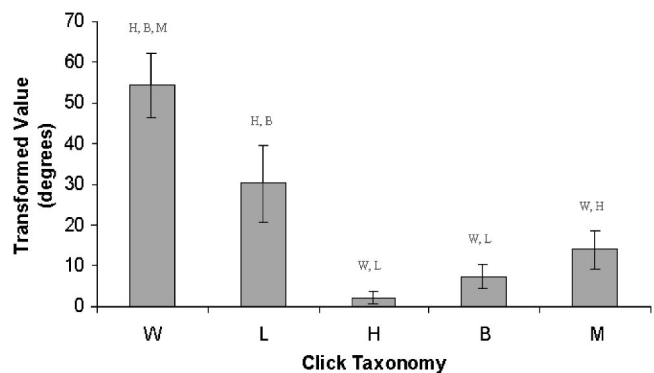


FIG. 3. Arc-sine transforms of the proportional use of specific click types for FLP. Statistically significant differences in the production of click types are denoted above a click category according to the following: H=high frequency, unimodal; L=low frequency, unimodal; B=bimodal; W=wideband; M=multimodal.

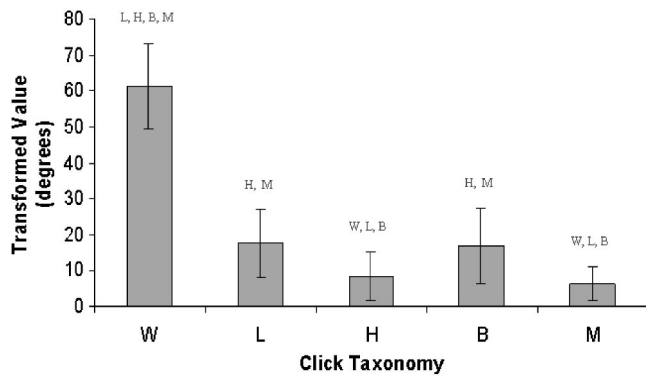


FIG. 4. Arc-sine transforms of the proportional use of specific click types for LUT. Statistically significant differences in the production of click types are denoted above a click category according to the following: H=high frequency, unimodal; L=low frequency, unimodal; B=bimodal; W=wideband; M=multimodal.

(Fig. 4, Kruskal–Wallis $SE=3722.6$, $q_{(0.05,\infty,5)}=3.858$). Wideband clicks were utilized to a significantly greater extent by LUT than any other type of click. Wideband clicks were utilized to a significantly greater extent by FLP for all but the low-frequency unimodal clicks. Transition between click bouts reflected a similar trend and indicated that high-frequency unimodal clicks were never transitioned to from low-frequency unimodal clicks (Tables III and IV; note that these tables only show results for target-absent trials but that similar results were also obtained on target-present trials).

The cumulative distribution of echolocation click peak frequencies was notably different between the two dolphins. Considering both target-present and target-absent trials, only 3.6% of all clicks produced by FLP contained peak frequencies greater than 60 kHz. In contrast, LUT produced clicks with peak frequencies above 60 kHz 20.4% of the time. The clicks of both animals demonstrated a bias towards particular groupings of peak frequencies; when plotted against the peak pressure level of the clicks, the groups clustered into bands. Clicks generated by LUT had a greater tendency to have peak frequencies of ~38, 54, and 69 kHz (Fig. 5). When peak frequencies occurred at higher frequencies, clustering was still notable but less defined. Peak frequencies of clicks emitted by FLP had a tendency to occur at 25, 35, and 40 kHz (Fig. 6). Although clicks produced by FLP occasionally had higher peak frequencies, a bias toward any specific peak frequencies above 60 kHz was not evident.

Raw detections, which provide an indication of how much clutter each animal encounters during its search, were overall greater for FLP than for LUT, though the number of clustered and raw detections experienced by LUT was more

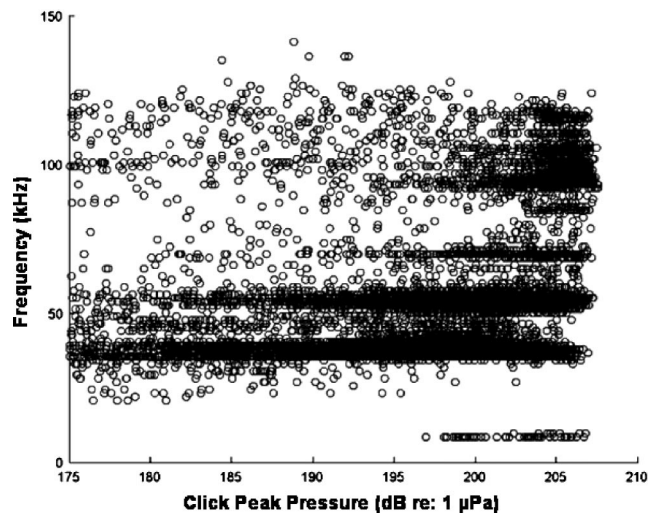


FIG. 5. Distribution of echolocation click-peak frequency as a function of click-peak pressure for LUT. Note the distinctive clustering around 38, 54, and 69 kHz.

variable (Fig. 7). Both animals experienced thousands of raw detections during target searches, and hundreds to in excess of a thousand detections when detections were range clustered within a 1-m range extent (i.e., consistent with multiple reflections from an object 1 m in diameter).

IV. DISCUSSION

Data collected from two dolphins, LUT and FLP, demonstrate differences in target detection strategies involving whole animal locomotion, the number of echolocation clicks produced, and the frequency content of those clicks. However, certain limitations to the data collection and methodology must be considered when interpreting the results.

First, statistical analysis of the instantaneous peak pressure of echolocation clicks is biased by the maximum level at which clicks could be recorded (peak sound-pressure level of 209 dB *re:1* μ Pa) and the threshold level accepted for analysis (peak sound-pressure level of 175 dB *re:1* μ Pa). Occasional clicks saturated the system and were excluded from analysis. Clicks produced below the threshold level likely occurred but were ignored as an acceptable trade-off between lost data and the minimization of detections due to noise. Consideration must also be given when comparing the analysis of echolocation clicks from target-present and target-absent trials, since the latter includes some amount of echolocation for the purpose of returning to the workboat from which the animal was sent. Furthermore, search and acquisition segments were defined by algorithmic detection

TABLE III. Transition probability matrix for click bouts according to their taxonomic classification for FLP during target-absent trials. Click taxonomy corresponds to the following: W (wideband), LFU (low frequency, unimodal), HFU (high frequency, unimodal), BI (bimodal), MULTI (multimodal).

	W	LFU	HFU	BI	MULTI
W		0.75	0.00	0.06	0.19
LFU	0.81		0.00	0.11	0.08
HFU	0.25	0.00		0.50	0.25
BI	0.30	0.36	0.05		0.29
MULTI	0.56	0.28	0.00	0.16	

TABLE IV. Transition probability matrix for click bouts according to their taxonomic classification for LUT during target-absent trials. Click taxonomy corresponds to the following: W (wideband), LFU (low frequency, unimodal), HFU (high frequency, unimodal), BI (bimodal), MULTI (multimodal).

	W	LFU	HFU	BI	MULTI
W		0.58	0.07	0.32	0.03
LFU	0.77		0.00	0.21	0.01
HFU	0.81	0.00		0.18	0.01
BI	0.63	0.29	0.04		0.04
MULTI	0.60	0.15	0.04	0.21	

and human verification of target echoes. This does not imply that the dolphin detection capability is equivalent; without a doubt the ability of the dolphin to detect and identify target echoes is superior to this approach. Finally, there are errors inherent in GPS-determined locations and in the onboard underwater navigation system of the BMT. Attempts to qualify and control these errors have been made through virtual reality playbacks and acoustic refinement of dolphin positions relative to targets. A certain degree of error is acceptable, particularly since movement patterns are only topically discussed here. However, this issue must eventually be addressed to permit more accurate localization of the dolphin and targets in three-dimensional space.

The search strategy of FLP can be grossly described as conservative. FLP demonstrated a stereotypical search pattern that consisted of descent to a point immediately above the benthos followed by a methodical search of the region around the search buoy. Click intervals and the number of clicks emitted were comparable between search and acquisition segments, and FLP swam within close proximity of the target prior to reporting it as present. Echolocation was stereotypical with search patterns typified by slow swim speeds and more distance traveled during positive trials (the animal typically swam around targets). In contrast, LUT was economical in click production, and on target-present trials traveled no more than a couple of meters from the boat prior to reporting target detection. This dolphin appeared to abide by the “law of least effort” in that energy spent to receive a

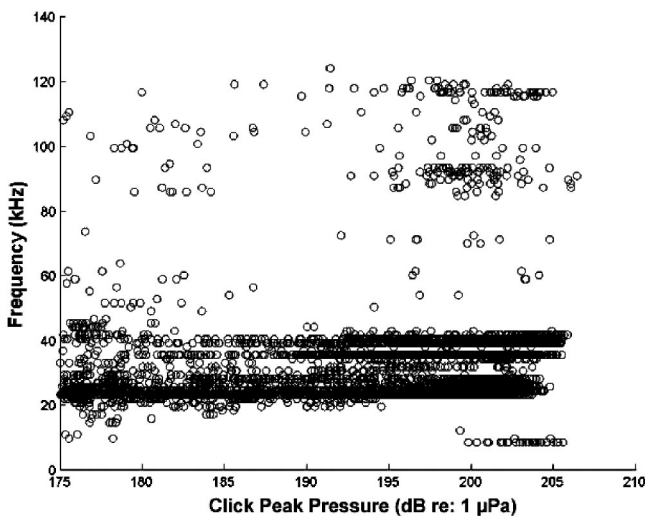


FIG. 6. Distribution of echolocation click-peak frequency as a function of click-peak pressure for FLP. Note the distinctive clustering around 25, 35, and 40 kHz.

food reward was minimized. Swimming around the buoy marker, 25–60 m away from the workboat, was only performed when no target was observed in the area and the dolphin had to complete the search pattern circuit to receive a reward. LUT echolocated more near the workboat at the start of the trial; relatively few clicks were produced in transit to the buoy, contrasting FLP in which copious numbers of clicks were produced throughout target present and absent trials.

Differences in the search strategies of the two dolphins are possibly explained by previous training histories and individual hearing ability. FLP was previously trained for open-water target searches in which targets, once found, were inspected again for verification prior to marking the identified target. In this study, reporting that a target is present only when in close proximity of the target may have resulted from integration of the previously acquired target marking behavior into this search strategy. In contrast, LUT was a naive animal with little experience in open-water target detection tasks. LUT may have developed a more economical strategy because no prior experience existed to im-

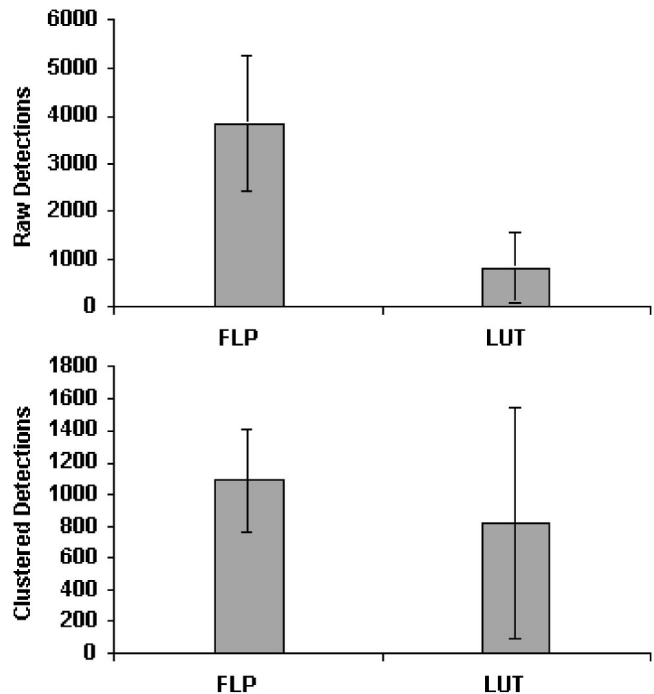


FIG. 7. Mean (\pm s.d.) of the number of detections received by FLP and LUT. Detections are presented as individual detections within the receive data stream, and detections clustered to a temporal run-out beyond the first detection equivalent to 1 m of travel time in the water (i.e., assuming a target of approximately 1-m diameter).

pact strategy development. Alternatively, recent hearing tests on FLP suggest that FLP has profound hearing loss above 50 kHz (Moore *et al.*, 2004). A reduction in hearing sensitivity may explain the reduced search range (shorter interclick intervals), predominant use of wideband clicks with low-frequency peaks, and more abundant click production. Similar characteristics have been observed in dolphins that have known hearing losses performing simpler target detection tasks (Houser *et al.*, 1999). Collectively, hearing ability and experience may have contributed to FLP evolving a conservative search strategy in which no false alarms were committed but correct detections were reduced.

Both LUT and FLP increased the source level of echolocation clicks during target acquisition segments. Although significant in both cases, the increase was smaller (~ 1 dB *re*: $1 \mu\text{Pa}$) for FLP than for LUT (~ 6 dB *re*: $1 \mu\text{Pa}$). Prior work by Au *et al.* (1988) suggests that the dolphin ear may act as an energy detector in the processing of echolocation clicks, and that multiple echoes of equivalent amplitude which do not fall within a critical time interval do not appear to impact detection thresholds for phantom echoes of similar frequency content. However, evidence from wild dolphins suggests that the pressure amplitude of emitted clicks may vary as $20 \log(R)$, where R is the distance to a target(s) that emits echoes as volume reverberation, as in a school of fish (Au and Benoit-Bird, 2003). Such variation is suggestive of time-varying gain controlled through source level variation. Unlike wild animals that vary click source level as a function of distance to a volume reverberant target, dolphins in this study searched for a target that is more like a point source and increased the echo level during acquisition of the target, prior to positive identification. Depending upon the subject, this was achieved by different means. In the case of FLP, an increase in the amplitude of echo returns was primarily achieved by decreasing the distance between the click source and the target. In contrast, LUT did not approach targets once detected, but increased the source level of his echolocation clicks prior to reporting detection. Although approached differently, these behaviors increase the amplitude of returning echoes and likely decrease the threshold for correct identification of a target through accentuation of echo highlights.

The tendency for echolocation clicks to have certain peak frequencies suggests either a preference for the use of certain frequencies or a resonant response of the click generating structures. The former possibility seems unlikely as most clicks produced by both animals were wideband, having -3 -dB bandwidths in excess of 85 kHz. However, separation between clusters of peak frequencies below 70 kHz is nearly harmonically related. Values for LUT are separated by ~ 15 kHz and the same is true for the lowest and highest values for FLP. Clustering at higher frequencies is evident, though difficult to quantify for LUT, and is absent for FLP. The separation in clusters is consistent with the idea of two sources for click generation, i.e., a pair of phonic lips (Cranford, 2000; Cranford *et al.*, 1996), if each set of lips has a different set of resonant frequencies resulting from its geometry and composition. This proposition is consistent with previous anatomical work that has demonstrated asymmetry in

the phonic lips of delphinid species (Cranford *et al.*, 1996). Furthermore, if FLP were constrained by hearing loss above 50 kHz, selective use of one set of phonic lips could allow the hearing loss to be accommodated by preferentially echolocating at frequencies at which there is functional hearing.

Both FLP and LUT demonstrated effective clutter rejection. The bathymetry over which the dolphins performed their searches was uneven, composed of heterogeneous sediment, and contained a number of clutter items, many of which were likely manmade. The ability of the animals to find and discriminate the search target via echoic means in the presence of many hundreds of reflective surfaces is a testament to their echolocation ability. The mechanism by which the dolphin excludes clutter is of interest to the sonar community, as replication of such capability would provide substantial improvement to current sonar detection and classification systems. It is uncertain whether the dolphins in this study learned the echo structure of the target and applied an exclusionary rule while searching (i.e., if not a , then reject), or whether an acoustic image of the total area searched is formed. Further research is required to answer such questions.

In summary, the use of remote acoustic and navigational recording devices carried by free-swimming dolphins in open water has provided new insight into dolphin echolocation search strategies. Data collected with new instrumentation (Martin *et al.* this issue) demonstrates that dolphins, with different hearing abilities and different training histories, are capable of finding targets in reverberant and cluttered environments but with a degree of success that varies as a function of distinct search strategies. The overwhelming occurrence of particular click-peak frequencies in the echolocation clicks of both dolphins is suggestive of structural resonances within two asymmetric click-producing structures. A lack of high-frequency peaks in one animal may indicate reliance upon a single set of structures for echolocation, the use of which may accommodate hearing loss above 50 kHz. Increasing the echo amplitude of targets during acquisition phases of a target search task was achieved by both dolphins through an increase in the source level of echolocation clicks and/or a reduction in the distance between target and dolphin without a reduction in click source level. Future experiments utilizing remote recording devices such as the BMT should provide an increasing body of knowledge on echolocation in freely swimming animals. By comparison to more traditional pen and pool work, a better understanding of how echolocation signals and strategies are varied in noisy, real-world environments will be obtained.

ACKNOWLEDGMENTS

The authors wish to acknowledge the financial support of the Office of Naval Research (Dr. Robert Gisiner, Dr. John Tague, Dr. Doug Abraham, and Dr. Mardi Hastings) and of the Defense Advanced Research Projects Agency (Dr. Alan Rudolph). Thanks goes to Robert Floyd of SPAWAR Systems Center San Diego for his consulting on measurements of dolphin echolocation parameters. For the training of diffi-

cult behaviors required in this project, the efforts of the SAIC contractor dolphin training crew are particularly appreciated.

- Altes, R. A., Dankiewicz, L. A., Moore, P. W., and Helweg, D. A. (2003). "Multiecho processing by an echolocating dolphin," *J. Acoust. Soc. Am.* **114**, 1155–1166.
- Altes, R. A., Moore, P. W., and Helweg, D. A. (1998). "Tomographic image reconstruction of MCM targets using synthetic dolphin signals," Technical Document: SPAWARSEN-San Diego, 17.
- Au, W., Popper, A. N., and Fay, R. R. (2000). "Hearing by Whales and Dolphins," in *Springer Handbook of Auditory Research*, edited by R. R. Fay and A. N. Popper (Springer, New York).
- Au, W. W. L. (1980). "Echolocation signals of the Atlantic bottlenosed dolphin (*Tursiops truncatus*) in open waters," in *Animal Sonar Systems*, edited by R. G. Busnel and J. F. Fish (Plenum, New York), pp. 251–282.
- Au, W. W. L., and Benoit-Bird, K. J. (2003). "Automatic gain control in the echolocation system of dolphins," *Nature (London)* **423**, 861–863.
- Au, W. W. L., Floyd, R. W., and Penner, R. H. (1974). "Measurement of echolocation signals in the Atlantic bottlenosed dolphin, *Tursiops truncatus*," *J. Acoust. Soc. Am.* **56**, 1280–1290.
- Au, W. W. L., and Moore, P. W. B. (1984). "Receiving beam patterns and directivity indices of the Atlantic bottlenose dolphin, *Tursiops truncatus*," *J. Acoust. Soc. Am.* **75**, 255–259.
- Au, W. W. L., Moore, P. W. B., and Pawloski, D. (1986). "Echolocation transmitting beam of the Atlantic bottlenose dolphin," *J. Acoust. Soc. Am.* **80**, 688–691.
- Au, W. W. L., Moore, P. W. B., and Pawloski, D. A. (1988). "Detection of complex echoes in noise by an echolocating dolphin," *J. Acoust. Soc. Am.* **83**, 662–668.
- Au, W. W. L., Pawloski, J. L., Nachtigall, P. E., Bonz, M., and Gisiner, R. C. (1995). "Echolocation signals and transmission beam pattern of a false killer whale (*Pseudorca crassidens*)," *J. Acoust. Soc. Am.* **98**, 51–59.
- Au, W. W. L., and Penner, R. H. (1981). "Target detection in noise by echolocating Atlantic bottlenose dolphins," *J. Acoust. Soc. Am.* **70**, 687–693.
- Brill, R. L., Moore, P. W. B., and Dankiewicz, L. A. (2001). "Assessment of dolphin (*Tursiops truncatus*) auditory sensitivity and hearing loss using jawphones," *J. Acoust. Soc. Am.* **109**, 1717–1722.
- Brown, M. B., and Forsythe, A. B. (1974). "Robust tests for the equality of variances," *J. Am. Stat. Assoc.* **69**, 364–367.
- Cranford, T. W. (2000). "In serach of impulse sound sources in odontocetes," in *Hearing by Whales and Dolphins*, edited by W. Au, A. N. Popper, and R. R. Fay (Springer, New York), pp. 109–156.
- Cranford, T. W., Amundin, M., and Norris, K. S. (1996). "Functional morphology and homology in the odontocete nasal complex—implications for sound generation," *J. Morphol.* **228**, 223–285.
- Freeman, M. F., and Tukey, J. W. (1950). "Transformations related to the angular and the square root," *Ann. Math. Stat.* **21**, 607–611.
- Helweg, D. A., Au, W. W. L., Roitblat, H. L., and Nachtigall, P. E. (1996). "Acoustic basis for recognition of aspect-dependent three-dimensional targets by an echolocating bottlenose dolphin," *J. Acoust. Soc. Am.* **99**, 2409–2420.
- Houser, D. S., Helweg, D. A., and Moore, P. W. B. (1999). "Classification of dolphin echolocation clicks by energy and frequency distributions," *J. Acoust. Soc. Am.* **106**, 1579–1585.
- Jacobs, D. W. (1972). "Auditory frequency discrimination in the Atlantic bottlenosed dolphin, *Tursiops truncatus* Montague: A preliminary report," *J. Acoust. Soc. Am.* **52**, 696–698.
- Johnson, C. S. (1968). "Relation between absolute threshold and duration-of-tone pulses in the bottlenosed porpoise," *J. Acoust. Soc. Am.* **43**, 757–763.
- Martin, S. W., Phillips, M., Bauer, E., Moore, P. W., and Houser, D. S. (2005). "Instrumenting free-swimming dolphins echolocating in open water," *J. Acoust. Soc. Am.* **117**, 2457–2463.
- Moore, P. W. B. (1997). "Mine hunting dolphins of the Navy," *SPIE—The International Society for Optical Engineering* **3079**, 2–6.
- Moore, P. W., Finneran, J., and Houser, D. (2004). "Hearing loss and echolocation signal change in dolphins," *J. Acoust. Soc. Am.* **116**, 2503.
- Roitblat, H. L., Au, W. W. L., Nachtigall, P. E., Shizumura, R., and Moons, G. (1995). "Sonar recognition of targets embedded in sediment," *Neural Networks* **8**, 1263–1273.
- Roitblat, H. L., Moore, P. W. B., Nachtigall, P. E., Penner, R. H., and Au, W. W. L. (1989). "Natural echolocation with an artificial neural network," *Int. J. Neural Networks* **1**, 239–248.
- Zar, J. H. (1984). "*Biostatistical Analysis*" (Prentice Hall, Englewood Cliffs, NJ).

Quantifying the acoustic repertoire of a population: The vocalizations of free-ranging bottlenose dolphins in Fiordland, New Zealand

Oliver Boisseau

Department of Marine Science, University of Otago, 310 Castle Street, Dunedin, New Zealand

(Received 1 June 2004; revised 28 December 2004; accepted 30 December 2004)

Quantification of the vocal repertoire of a species is critical for subsequent analysis of signal functionality, geographical variation, social relevance, and transmission. While signal repertoires have been documented for numerous animal species, detailed descriptions for the cetaceans are rare. This study describes the vocalizations of free-ranging bottlenose dolphins resident to the waters of Fiordland, New Zealand. The field recordings presented in this paper were made using both audio band and broadband apparatus. Subsequent classification of vocalizations using multivariate parameters enabled a quantification of the entire vocal output of the focal animals. These results were used to propose a meaningful repertoire of signals employed by this species. A total of 12 individual signal types were described, comprising four broad structural classes: “tonal,” “single bursts,” “click bursts,” and “repeat bursts.” The proposed repertoire will allow subsequent investigation into vocal behavior. It appears that the successful description of a species’ repertoire is dependent on the use of appropriate recording systems, a high number of representative recordings with good signal-to-noise ratio, and subsequent validation of the original classification system. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1861692]

PACS numbers: 43.80.Ka [WA]

Pages: 2318–2329

I. INTRODUCTION

Quantifying the signals produced by a species is a fundamental first step in the investigation of their functionality, geographical variation, social relevance, and cultural transmission. Signaling acts may include physical displays, pheromone release, and vocalizations (Herman and Tavolga, 1980). The vocal repertoires of numerous vertebrate species have been described, including such diverse taxa as the anurans (Grafe *et al.*, 2000), songbirds (Kroodsma and Henri, 1982), ungulates (Budde and Klump, 2003), carnivores (Volodina, 2000), and primates (Fischer and Hammer-schmidt, 2002). The cetaceans comprise an extremely vocal mammalian order [see Herzing (2000) for a recent review] and numerous aspects of their sound production and auditory systems have been described in detail. Surprisingly, thorough descriptions of vocal repertoires are rare amongst the cetacea. For the baleen whales, for example, a comprehensive repertoire description has been provided for a population of southern right whales *Eubalaena australis* (Clark, 1982) including vocal and nonvocal calls. A clustering technique has been applied to the signals of humpback whales *Megaptera novaeangliae* to derive an objective classification of calls (Chabot, 1988).

Amongst the toothed whales, perhaps the best documented communication system is that of the beluga *Delphinapterus leucas*, in which 80% of calls may be pulsive (Sjare and Smith, 1986; Recchia, 1994). Click-based repertoires from two other odontocete species have also been documented: the Hector’s dolphin *Cephalorhynchus hectori* (Dawson and Thorpe, 1990) and the sperm whale *Physeter macrocephalus* (Gordon, 1987). The repertoire of resident killer whales *Orcinus orca* has been described in some detail and is

based principally upon clicks, pulsed calls, and whistles (Ford, 1989). Vocalizations from false killer whales *Pseudorca crassidens* have been described along a continuum with clicks and whistles representing extreme cases (Murray *et al.*, 1998a). Spinner dolphin *Stenella longirostris* calls have been categorized as burst pulses or whistles, with some further division within these categories (Brownlee and Norris, 1994). Calls produced by Risso’s dolphins *Grampus griseus* have been divided into seven categories (Corkeron and Van Parijs, 2001), the calls of Irrawaddy dolphins *Orcaella brevirostris* into six (Van Parijs and Corkeron, 2001), and calls of the tucuxi *Sotalia fluviatilis* into four (Monteiro and Monteiro, 2001). A number of other studies (e.g., Lilly and Miller, 1961; Caldwell and Caldwell, 1967; Ridgway, 1983) have classified vocalizations under the three broad structural classes of whistles (narrow-band, tonal signals), clicks (broadband pulses with rapid onset/offset), and burst pulses (sequences of impulsive sounds).

Almost all of the vocal repertoires proposed for the cetaceans are based on classification of aural and spectrographic properties. While useful in reducing the complexity of a dataset, there is often no subsequent testing of the validity of the proposed categories. Notable exceptions include the work on right whales (Clark, 1982), humpback whales (Chabot, 1988), Hector’s dolphins (Dawson and Thorpe, 1990), and beluga (Recchia, 1994) and false killer whales (Murray *et al.*, 1998b).

Another pitfall in the categorization of signaling systems is the equipment used for recordings. Many cetaceans produce vocalizations within the human audio range (below 20 kHz) and even solely ultrasonic species may be audible at lower frequencies through an artifact of click repetition rate (Watkins *et al.*, 1977). However, significant energy may ex-

TABLE I. Synopsis of signal types proposed in this study in relation to other studies of bottlenose repertoire. Context for signals inferred from surface behavior (Boisseau, 2005). Signal prevalence is the percentage of all 875 recordings that contain each signal type.

Signal type	Prevalence		Synonym	Study
	(%)	Fiordland context		
Buzz	60	Foraging	Buzz	Lilly and Miller (1961); Jacobs <i>et al.</i> , (1993); Herzing (1996)
Choke	19	Cavitation	Mew	Wood (1953); Goodson <i>et al.</i> (1988)
			Bark	Wood (1953)
			Chuff	McCowan and Reiss (1995)
Click trains	66	Echolocation	Kerplunk	Connor <i>et al.</i> (2003)
			Click trains	Summaries in Richardson <i>et al.</i> (1995) and Herzing (2000)
Cough	30	Arousal	Gulp	dos Santos <i>et al.</i> (1995)
Creak	57	Foraging	Grunt	Hastie (2003)
			Grate	Wood (1953)
Hiccup	43	Arousal	Creak	Lilly and Miller (1961); dos Santos <i>et al.</i> (1990)
			Ranging clicks	Goodson <i>et al.</i> (1988)
			Gulp	dos Santos <i>et al.</i> (1995)
Laughter	24	Agonistic	Thunk	McCowan and Reiss (1995)
Long styro	4		Bray	dos Santos <i>et al.</i> (1995); Hastie (2003); Janik (2000b)
Orca	31	Agonistic	Screech	dos Santos <i>et al.</i> (1995)
			Squawk	Caldwell and Caldwell (1967); Lilly and Miller (1961)
Quack	32	Arousal	Yelp	Wood (1953)
			Quack	Lilly and Miller, (1961); Jacobs <i>et al.</i> (1993); Hastie (2003)
Ratchet	39	Agonistic	Knocking	Gish (1979); Connor and Smolker (1996)
Short styro	33	Agonistic	Pop	Connor and Smolker (1996)
			Bray	See above
Splash	35	Water contact	Crack	Caldwell and Caldwell (1967)
			Bang	dos Santos <i>et al.</i> (1995)
Squeal	27	Conciliation	Jaw pop	Finneran <i>et al.</i> (2000)
			Squeak	Caldwell and Caldwell (1967)
			Whistle-squawk	Reiss (1988)
Whistle	77	Contact call	Chirp	Hastie (2003)
			Whistle	Summaries in Richardson <i>et al.</i> (1995) and Herzing (2000)

tend into higher frequency ranges (i.e., above 20 kHz). The equipment required for recording these higher frequencies is often expensive and/or cumbersome. As such, only investigations into echolocation clicks have tended to use broadband systems [see Herzing (2000) for review]. Quantitative broadband analyses of vocalizations that are not solely echolocatory are restricted to the Hector's dolphin (Dawson, 1991), spinner dolphin (Brownlee and Norris, 1994), spotted dolphin (Lammers *et al.*, 2003), and white-beaked dolphin (Rasmussen and Miller, 2002).

Bottlenose dolphins (*Tursiops truncatus*) are extremely vocal and numerous studies have been made of their acoustic behavior. The majority of these studies have concerned the echolocation system in captive contexts (see Au, 1993 for a review). Of the studies based upon free-ranging animals, most emphasis has been placed on whistles. Whistles are tonal signals that appear to be used as contact calls between individuals in some situations (e.g., Sayigh *et al.*, 1999; Smolker and Pepper, 1999; Janik, 2000a; Jones and Sayigh, 2002). These vocalizations may be recorded readily in the field in high numbers with good signal-to-noise ratios. The presence of numerous other signals within the repertoire of

bottlenose dolphins has long been acknowledged and many appear universal (some are summarized in Table I). However, these vocalizations have received much less attention for various reasons. Some are produced far less frequently than whistles and clicks, and thus require high levels of study effort to build up large samples. These signals may be produced with low amplitude levels, and hence must be recorded from animals at close quarters. Any overlapping or intergrading of signals may create an obstacle for meaningful classification. A broad range of physical descriptions has been applied to pulsed bottlenose vocalizations, yet many of these descriptions may refer to fundamentally similar sounds. Many of these classifications are subjective, based on differences perceived by human observers. They are thus unquantified and mostly unrepeatabe.

Amid the abundant literature pertaining to bottlenose dolphins, there are no descriptions of complete vocal repertoires for closed populations, either in captivity or the wild. Descriptions of various signal types exist (e.g., Connor and Smolker, 1996; dos Santos *et al.*, 1995; Herzing, 1996; McCowan and Reiss, 1995; Schultz *et al.*, 1995), but little mention has been made of the individuality or overlap of these

vocalizations. The most comprehensive review of bottlenose vocal activity is that of Caldwell and Caldwell (1967) and was largely based on two small captive groups. Likewise, there is a dearth of information on the broadband properties of the non-echolocatory signals of bottlenose dolphins. Broadband results from captive studies have been published for one example of a whistle (Lilly and Miller, 1961), as well as 26 cases of a “jaw pop” vocalization and 17 cases of nonvocal breaches and tail slaps (Finneran *et al.*, 2000). For free-ranging dolphins, an extremely brief mention has been made of broadband recordings of “pop” trains (Connor and Smolker, 1996).

The waters of Fiordland in the southwest of New Zealand’s South Island are home to three separate resident populations of bottlenose dolphin. Data from two of these populations are presented in this paper. The fiords present ideal study sites for work on free-ranging dolphins. The steep rock walls create shelter from prevailing weather patterns. The dolphins tend to follow predictable seasonal patterns and these waterways effectively provide conditions of “natural captivity.” Individuals from Doubtful Sound rarely leave the fiord (Schneider, 1999) while the home range of the Milford Sound population encompasses seven northern fiords, several bays, and one lake system (Dave Rundgren, personal communication). The aims of this paper are (i) to compare vocalizations of free-ranging bottlenose dolphins using both audio band and broadband apparatus, (ii) to quantify the entire vocal output of the focal animals, (iii) to classify the vocalizations using multivariate parameters, and (iv) to propose a meaningful repertoire for subsequent investigations in to vocal behavior.

II. METHODS

Recordings of bottlenose dolphins were made between August 2000 and June 2003 from two resident populations in Fiordland, southwest New Zealand. Surveys for dolphins were conducted from a 4.8-m aluminum research vessel, with a 50-hp Yamaha 4-stroke engine. Acoustic recordings were made while drifting ahead of dolphin groups with the engine off. Recordings were made using a Sonatech 8178 mono element hydrophone (30 Hz to 180 kHz \pm 3 dB) on 10 m of cable connected to an on-board preamplifier with a linear frequency response (dc—300 kHz). For recordings made within the human audio range, a Sony TCD D8 digital audio tape recorder was used with a calibrated frequency response of 10 Hz to 22 kHz \pm 1 dB (sampling at 48 kHz). For broadband recordings, an analog Racal Store 4DS recorder was used at a speed of 30 ips, for which the calibrated frequency response is 120 Hz to 150 kHz \pm 0.5 dB. For lower frequencies the DAT recorders are far superior to the Racal, offering a broader dynamic range (>85 vs. 40 dB) and excellent fidelity to infrasonic frequencies (10 Hz).

A. Tape transcription

All DAT and Racal recordings made in the field were reviewed at a later date to avoid observer bias. Acoustic emissions were transcribed by noting the time and the type of signal. All recorded vocalizations belonged to one of four structural classes: “tonal” (narrow-band, frequency-

modulated signals), “single burst” (short impulsive emissions with most energy below 5 kHz), “click burst” (broadband pulses with rapid onset/offset time), and “repeat burst” (sequences of single bursts). Within these classes, 15 signal types were initially transcribed: *buzz*, *click train*, *creak*, *choke*, *cough*, *hiccup*, *laughter*, *long styro*, *orca*, *quack*, *ratchet*, *short styro*, *splash*, *squeal*, and *whistle*. Signal categorization was based in part on previously reported vocalizations and partly on novel interpretation. Although qualitative categorization of this kind is extremely subjective, it allows a division and reduction of the data for later analysis. Efforts were made to transcribe every individual vocalization heard, with the exception of clicks. As echolocation clicks were present in almost all recordings, often in overwhelming numbers, it was not possible to define individual click trains.

B. Digitization of vocalizations

For subsequent analysis, vocalizations were randomly sampled from all seasons for all three years of data to control for seasonal variation. The target sample size was arbitrarily set at 10 signals per season per year for each proposed signal type for both Doubtful and Milford Sounds. Thus for any given signal type, a maximum of 10 (samples) \times 4 (seasons) \times 3 (years) \times 2 (fiords) = 240 samples would be available. To limit repeated sampling of individual dolphins, selected samples were at least a day apart when possible. As higher frequency signals may be highly directional, only vocalizations with relatively high amplitude were digitized in an effort to enhance the likelihood of measuring on-axis signals. Vocalizations were deemed suitable for analysis if their mean square (rms) amplitude was more than 10 dB above background noise. Vocalizations were digitized from DAT tapes with a sample rate of 44.1 kHz via a 24 bit RME Digi96/8 digital sound card. To avoid aliasing, signals were filtered using a Kemo VBF8 filter set at 22 kHz with 96 dB/oct attenuation. Wideband Racal recordings were digitized at an effective sampling rate of 352.8 kHz ($\frac{1}{8}$ record speed at 44.1 kHz; 16 bit). Measurements were taken from digitized vocalizations using Canary 1.2.4. (Cornell Laboratory) on a 933-MHz G4 Macintosh computer.

C. Comparison of broadband recordings with DAT recordings

Of the 875 recordings made in both fiords, 25 were made using the broadband system during periods of high vocal activity. From these recordings, all vocalizations with suitable signal-to-noise ratio were digitized and measured using Canary’s automatic measurement function. Peak frequencies were measured and compared to similar values from DAT recordings using a Student’s *t*-test. All measurements were log-transformed to approximate the normal distribution. As the sample size for the Racal system was markedly smaller than for the DAT system, vocalizations were only tested if a suitable sample size was available using the following iterative rule:

$$n = \frac{s^2 t_{\alpha(2), (n-1)}^2}{d^2},$$

where s^2 is the sample variance, t is the two-tailed critical value of Student's statistic, d is the half width of the desired confidence interval, and $1 - \alpha$ is the confidence level for the confidence interval (Zar, 1998). In addition, nine long recordings were made on separate occasions using both systems simultaneously. The spectrograms from these "duets" were compared for evidence of ultrasonic components.

D. Quantitative measurement of signal parameters

Signal parameters were measured from waveform, spectrogram, and point spectrum windows. Spectrographic information was obtained using fast Fourier transformation (FFT) with a Hamming window function. The FFT procedure has an inherent trade-off between time resolution and frequency resolution (Bradbury and Vehrencamp, 1998). As a wide variety of signal types were measured in this study, a window size of either 256 or 512 points was selected on a file by file basis to allow the measurement of time-frequency parameters from each vocalization. The parameters measured from each vocalization were dictated by its structure. However, from the diverse range of signals recorded, six parameters could be measured from all vocalizations. The first of these, "peak frequency," was defined as the frequency at which a signal has the highest energy. The time of this peak intensity ("peak time") was also measured along with "center time" (the point at which the collective intensity of a signal is half the total). Both these values were measured as proportions of "total signal duration." Interpeak interval ("IPI") was measured as the time between successive pressure peaks in a vocalization. "Duty cycle" was measured for each vocalization and is defined as the percentage of time a signal is "on" in relation to its total length (Murray *et al.*, 1998a).

E. Spectrographic cross-correlation

The spectrographic cross-correlation (SPCC) procedure performs quantitative comparisons between spectrograms by incrementally sliding spectrograms past each other and calculating correlation coefficients at successive time offsets. Thus SPCC can provide a quantitative measure of vocal similarity that is both automated and repeatable. As the similarity is being measured between spectrograms that may be radically different in length, peak correlation is of interest rather than time offset. The SPCC procedure was employed using Canary's batch correlator function. Start and end times of individual signals were determined by close inspection of spectrograms and waveforms. During SPCC evaluation, the signals being compared were normalized to ensure the maximum amplitudes of both signals were the same. Spectrograms were derived using the same settings (FFT and frame length 512 points; filter bandwidth 350 Hz; 75% overlap; Hamming function). Each spectrogram in the dataset was compared with every other spectrogram in turn. This process provided a measure not only of the similarity between individual signals, but also between signal types (*whistle* versus *buzz*, etc.). Hierarchical cluster analysis was conducted on the resulting distance matrix. Ward's linkage method was used to amalgamate vocalizations based on increasingly dis-

similar SPCC values. This method tends to produce compact groups (Ward, 1963).

F. Principal component analysis of signal type

Principal component analysis (PCA) is of use when *reducing* the number of variables (by removing highly correlated variables) and *detecting structure* between variables. The principal components method finds linear combinations of variables (components) that account for as much variation in the original variables as possible. Usually, a few components will account for most of the variation, and these components can be used in place of the original variables during subsequent analysis (Stevens, 2001). Principal components were extracted using a variance maximizing (varimax) rotation that maintains the cumulative percentage of variation explained by the extracted components but spreads it more evenly over the components. Thus the rotated component matrix is easier to interpret than the unrotated matrix. The PCA scores for each vocalization were used in discriminant function analysis (DFA) to investigate differences between signal types. DFA is used as a descriptive technique to determine which variables discriminate between two or more naturally occurring groups. The accuracy of the DFA model was tested using cross validation, a process that classifies each case while leaving it out from the model calculations.

III. RESULTS

Over the 3-year study period, 875 recordings were made of bottlenose dolphins in Fiordland. From 49 hours of recorded vocal output, approximately 50 000 separate vocalizations were categorized aurally into four broad structural classes—"tonal," "single burst," "click burst," and "repeat burst." These classes incorporated 15 discrete signal types. Two of the signal types (the single burst *laughter* of Doubtful Sound and the triple-burst *long styro* of Milford Sound) were fiord specific. The calls in the proposed repertoire were also segregated to a certain extent by behavioral context (Boisseau, 2005).

A. Description of signal types

Mean spectral and structural qualities were calculated for each of the proposed signal types (Table II). For the majority of vocalizations recorded, most energy was below 22 kHz (i.e., the upper frequency limit of the DAT system). However, certain vocalizations had significant ultrasonic energy, typically the click-based signals and the harmonics of *whistles*. Some vocalizations, such as *laughter* and the *styro*s, seem to possess low-frequency narrow-band components, which (in the case of the *styro*s) oscillated rhythmically with broadband pulses.

"Tonal" signals are comprised of two separate types. *Whistles* are narrow-band frequency-modulated signals that are typically longer than 500 ms and have most energy between 7 and 26 kHz. *Squeals* are also narrow-band signals

TABLE II. Summary of the properties of all 15 proposed signal types. Values are means ($\pm 95\%$ confidence limits). Duty cycle generally increases moving down the table. The occurrence of each signal type is measured as a percentage of all 49 222 transcribed calls.

	% of all calls	Total duration (ms)	Peak frequency (Hz)	IPI (ms)	Duty cycle	Max. harmonic interval	Max. frequency (kHz)	No. of repeated elements	Burst length (ms)	Inter-burst interval (ms)	Interrepeat interval (ms)
Tonal											
Whistle	33.4	857 (± 79)	10401 (± 364)	0.08 (± 0.00)	1.00 (± 0.0)	15577 (± 562)	41 (± 6)				
Squeal	1.7	181 (± 39)	2914 (± 578)	0.32 (± 0.06)	1.00 (± 0.0)	2513 (± 499)	13 (± 9)				
Single bursts											
Splash	3.4	42 (± 6)	1269 (± 236)	0.33 (± 0.03)	1.00 (± 0.0)	639 (± 79)	115 (± 19)				
Choke	1.3	135 (± 19)	1281 (± 265)	0.27 (± 0.03)	1.00 (± 0.0)	658 (± 104)	109 (± 53)				
Cough	1.3	120 (± 15)	692 (± 160)	1.02 (± 0.19)	1.00 (± 0.0)	541 (± 138)	22 (± 8)				
Quack	2.4	147 (± 19)	1250 (± 337)	0.68 (± 0.09)	1.00 (± 0.0)	476 (± 25)	66 (± 56)				
Click bursts											
Orca	4.4	316 (± 36)	1863 (± 476)	0.46 (± 0.06)	0.63 (± 0.05)	552 (± 41)	90 (± 33)	62 (± 7)	2.5 (± 0.2)	3.0 (± 0.3)	
Buzz	22.4	1004 (± 143)	2389 (± 484)	0.33 (± 0.05)	0.38 (± 0.02)	478 (± 20)	78 (± 21)	167 (± 24)	0.3 (± 0.1)	3.9 (± 2.0)	
Creak	13.6	1068 (± 142)	1494 (± 334)	0.38 (± 0.07)	0.24 (± 0.06)	502 (± 47)	137 (± 28)	73 (± 38)	4.2 (± 0.2)	30.2 (± 4.3)	
Ratchet	6.1	1132 (± 158)	1353 (± 397)	0.74 (± 0.08)	0.22 (± 0.05)	499 (± 38)	89 (± 42)	14 (± 2)	16.7 (± 2.0)	77.3 (± 3.8)	
Click train	a	3002 (± 762)	3356 (± 1780)	0.18 (± 0.06)	0.08 (± 0.04)	689 (± 114)	118 (± 25)	54 (± 8)	4.2 (± 0.8)	90.6 (± 37.8)	
Repeat bursts											
Hiccup	3.9	312 (± 30)	805 (± 145)	1.33 (± 0.17)	0.54 (± 0.06)	483 (± 31)	70 (± 31)	2 (± 0)	134.7 (± 28.1)	3.4 (± 0.4)	154 (± 22)
Laughter	2.6	4198 (± 1131)	1058 (± 416)	1.55 (± 0.20)	0.23 (± 0.05)	495 (± 48)	11 (± 5)	15 (± 4)	74.6 (± 10.4)	4.6 (± 0.6)	266 (± 32)
Short styro	2.4	4101 (± 584)	1071 (± 312)	1.13 (± 0.13)	0.38 (± 0.08)	496 (± 33)	21 (± 5)	6 (± 1)	73.5 (± 35.1)	2.8 (± 0.2)	372 (± 40)
Long styro	0.3	8159 (± 1551)	1079 (± 280)	0.88 (± 0.20)	0.51 (± 1.00)	564 (± 85)	35 (\dots)	7 (± 2)	141.0 (± 17.8)	2.4 (± 0.7)	221 (± 60)

^aCounts of individual *click trains* are excluded due to the problems associated with identifying discrete trains.

but differ from *whistles* in that they are shorter (usually less than 300 ms) and of lower frequency (most energy below 5 kHz). *Squeals* also tend to be upsweeps (i.e., the fundamental frequency increases).

The rest of the proposed repertoire is pulse based. Vocalizations in the “single burst” class are intrinsically short. *Splash* is the shortest vocalization in the proposed repertoire (less than 50 ms) and is composed of a single large pulse with rapid onset and offset. *Choke* is structurally similar to *splash* yet has an additional noisy broadband element that follows the initial pulse and thus increases signal duration (~ 135 ms). *Cough* and *quack* are of similar duration (~ 135 ms), but *cough* is generally a noisy signal with very low fundamental frequency (700 Hz) while *quack* is almost sinusoidal in nature, with higher fundamental frequency (1250 Hz).

The remaining pulsed vocalizations consist of repeated bursts of pulses. The “click bursts” are composed of regular broadband clicks and are structurally similar but differ mostly in interburst interval (IBI), i.e., the time between successive clicks. *Click trains* have the highest IBI (90 ms) and

also tend to be relatively long (upwards of 2000 ms). *Ratchets* have lower IBIs (75 ms) and are composed of clicks with strong emphases on lower frequencies. *Creaks* have still lower IBI values (30 ms) and tend to be broadband in nature. *Buzzes* have such low IBI values (4 ms) that consecutive clicks start to overlap and results in a single tone to the human ear. *Orcas* have the lowest IBI of the click bursts (3 ms) and also sound tonal to a human listener. *Orcas* tend to be shorter than *buzzes* (320 and 1000 ms, respectively). In many ways *orcas* may be seen as intermediaries between the click bursts and the single bursts.

The final structural subclass of the pulsed vocalizations is the “repeat bursts.” These signals are low in frequency (typically below 2 kHz) and have repeated elements, but do not possess the broadband nature of click bursts. The simplest of these repeat bursts is the *hiccup*, a short vocalization (310 ms) composed of two separate burst trains (essentially a double *cough*). *Laughter* is similar in structure to *hiccup* but is composed of numerous *coughlike* bursts (typically 15, although examples with over 120 repeats have been recorded). *Short styro* is formed of repetitive cycles (usually six) and

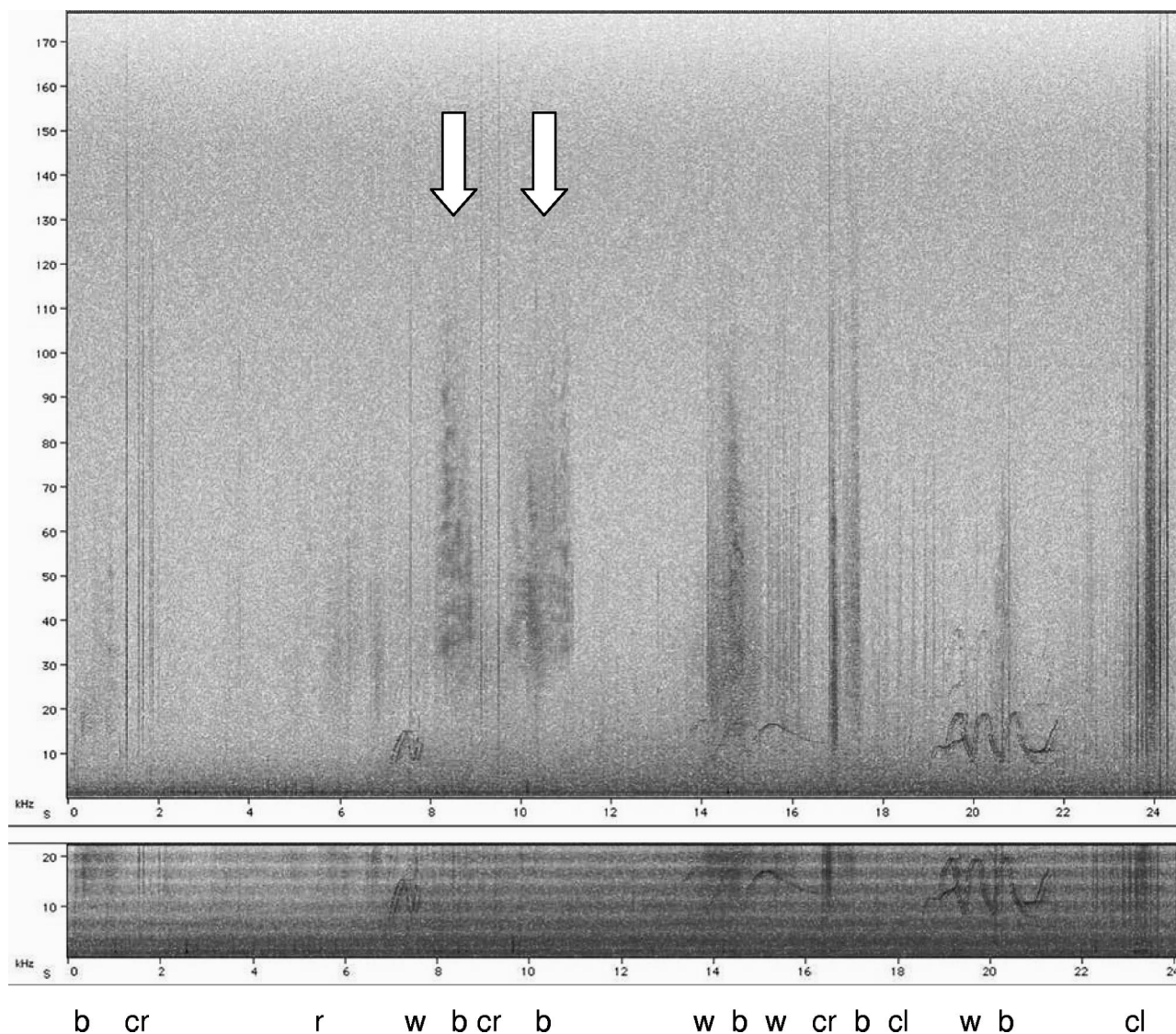


FIG. 1. Example of consecutive “duet” recording (broadband Racal spectrogram above narrowband DAT spectrogram). Arrows denote two solely ultrasonic *buzzes*. Note boat noise apparent during the DAT recording but absent from the RACAL spectrogram due to differences in dynamic range. Individual vocalizations are represented as b=*buzz*, cl=*click train*, cr=*creak*, r=*ratchet*, and w=*whistle*.

each cycle is composed of two variable burst trains. *Long styro* is similar although each cycle is formed by three variable burst trains. As the number of repeated elements forming *laughter* and the *styros* is highly variable, the durations of these sequenced vocalizations are also highly variable (from 100 ms to 120 s).

B. Comparison of broadband recordings with DAT recordings

Few of the proposed signal types exhibited a greater emphasis in the ultrasonic range than the audible range. The only vocalizations that seemed to contain minimal energy in the DAT range (i.e., below 24 kHz) were *click trains*. Comparisons between consecutive DAT and Racal recordings suggest only *click trains* and *buzzes* may ever be exclusively ultrasonic. For example, the sequence portrayed in Fig. 1 comes from a longer recording containing 110 separate vocalizations of which only 2 are ultrasonic (both *buzzes*). This

is in keeping with the other eight “duet” recordings, for which solely ultrasonic vocalizations typically represent less than 2% of all signals recorded.

Of the parameters measured from each vocalization, peak frequency is the measurement most likely to be underestimated by DAT recordings. To investigate this misrepresentation, the peak frequencies of vocalizations were measured from both DAT and Racal recordings. The lower frequency limit of the DAT system is below that of the Racal system (30 and 120 Hz, respectively, using a Sonatech hydrophone). As approximately 3% of peak frequencies measured from the DAT tapes were below 120 Hz, these values were excluded from analysis. Log transformed values from Racal recordings were significantly higher than DAT recordings for *buzz*, *hiccup*, *quack*, *short styro*, *splash*, *squeal*, and *click trains* (Student *t*-test; all $p < 0.01$). However, all mean values were below 22 kHz (the upper DAT limit) except for the *click trains* (Fig. 2). Only in the click-based categories (*click trains*, *creak*, *buzz*, and *orca*) were peak frequencies

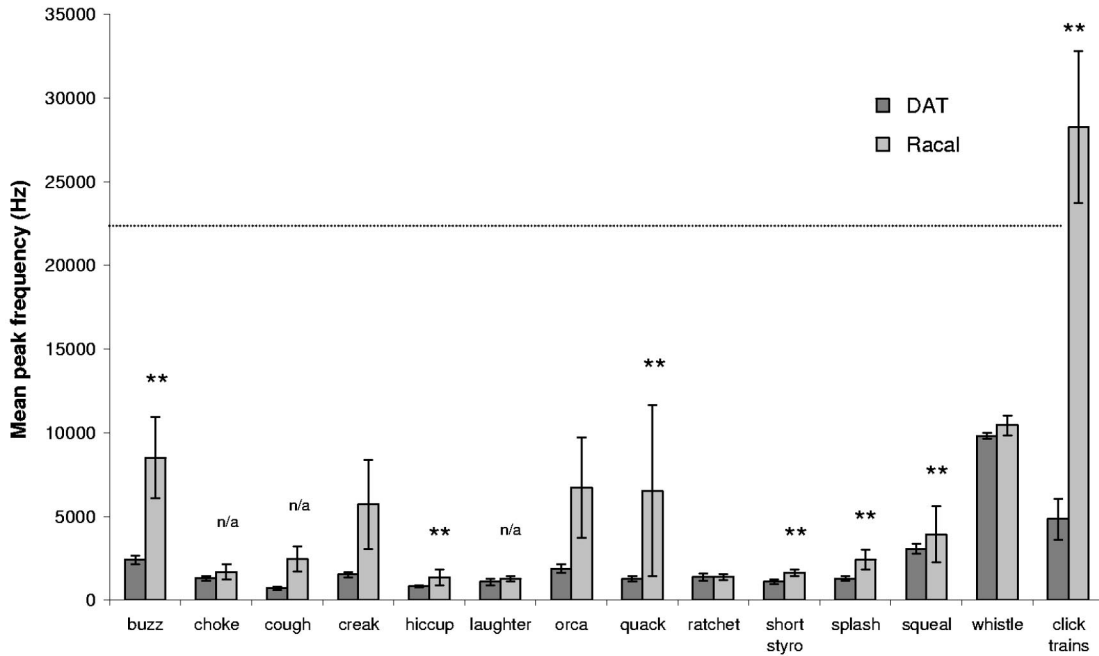


FIG. 2. Summary of mean peak frequencies for all proposed signal types. Error bars represent standard errors of the mean. Dotted line represents upper frequency limit of DAT system. *Long styro* excluded due to low number of broadband samples. (**) represents significant *t*-test differences at the 99% level. (n/a) represents those signals with too few RACALS samples to conduct Student's *t*-test.

measured above 22 kHz. Recordings made with the DAT system seem to be representative of all signal types except broadband *click trains*. It must be noted that the sample size of Racal recordings is low and thus the *t*-tests have relatively low statistical power.

C. Quantitative measurement of signal parameters

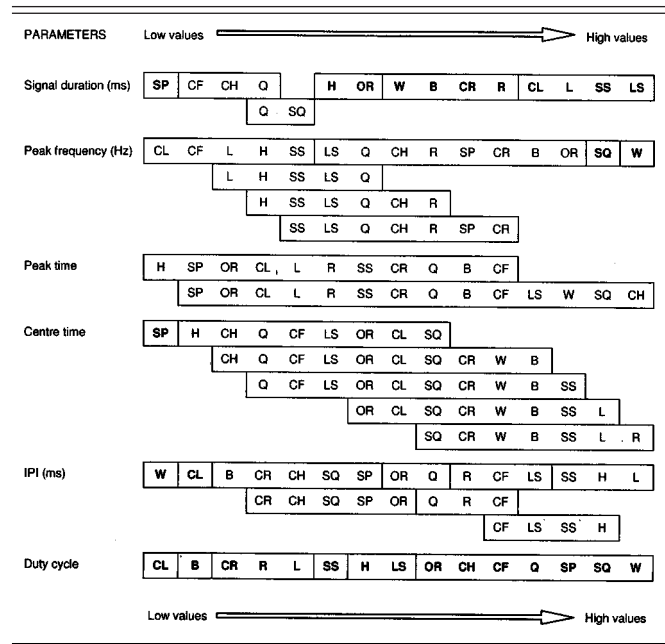
A dataset of 1959 individual vocalizations was extracted from the 49 h of recordings. These samples represented 3 years of data from both fiords, incorporating all 12 seasons in that period. To compare signal types a one-factor ANOVA was conducted on the log-transformed data. For each of the six common parameters, significant differences existed between the 15 proposed signal types (all $p < 0.01$). High *F* values (> 100) for “signal duration,” “peak frequency,” and “duty cycle” suggest variation in these parameters is much greater between groups than within groups. Tukey's range test was subsequently used to identify homogenous subsets of signal types (Table III).

D. Spectrographic cross-correlation

Spectrograms were produced for all 1959 vocalizations in the dataset and were subjected to spectrographic cross-correlation (SPCC). When comparing signal types, the mean peak correlation values were lowest for *splash*, *long styro*, *short styro*, and *click trains*, suggesting the spectrograms of these vocalizations were relatively distinct. A dissimilarity matrix was derived from the SPCC matrix and a dendrogram formed using hierarchical cluster analysis (Fig. 3). The dendrogram reveals two main clusters that may broadly be described as repetitive (composed of repeated bursts) and non-repetitive (composed of single bursts). Notable exceptions to these clusters are *whistle*, *orca*, and *hiccup*. *Whistle* forms a

cluster with some of the click-based signals (*buzz*, *creak*, and *ratchet*). These click-based vocalizations all have energy covering a broad range of frequencies. As the peak frequencies of *whistles* tend to be higher than most other vocalizations (Fig. 2), it is likely that any given *whistle* contour will have a high correlation with broadband vocalizations. *Orca* and *hiccup* are both repetitive vocalizations, *orca* being composed of rapid clicks and *hiccup* being composed of twin

TABLE III. Summary of Tukey's *posthoc* range tests for one-factor ANOVA. Vocalizations contained within the same rectangle are not significantly different from each other at the 95% level. Rectangles that do not overlap (in bold) are significantly distinct from all others. See Fig. 4 for vocal abbreviation key.



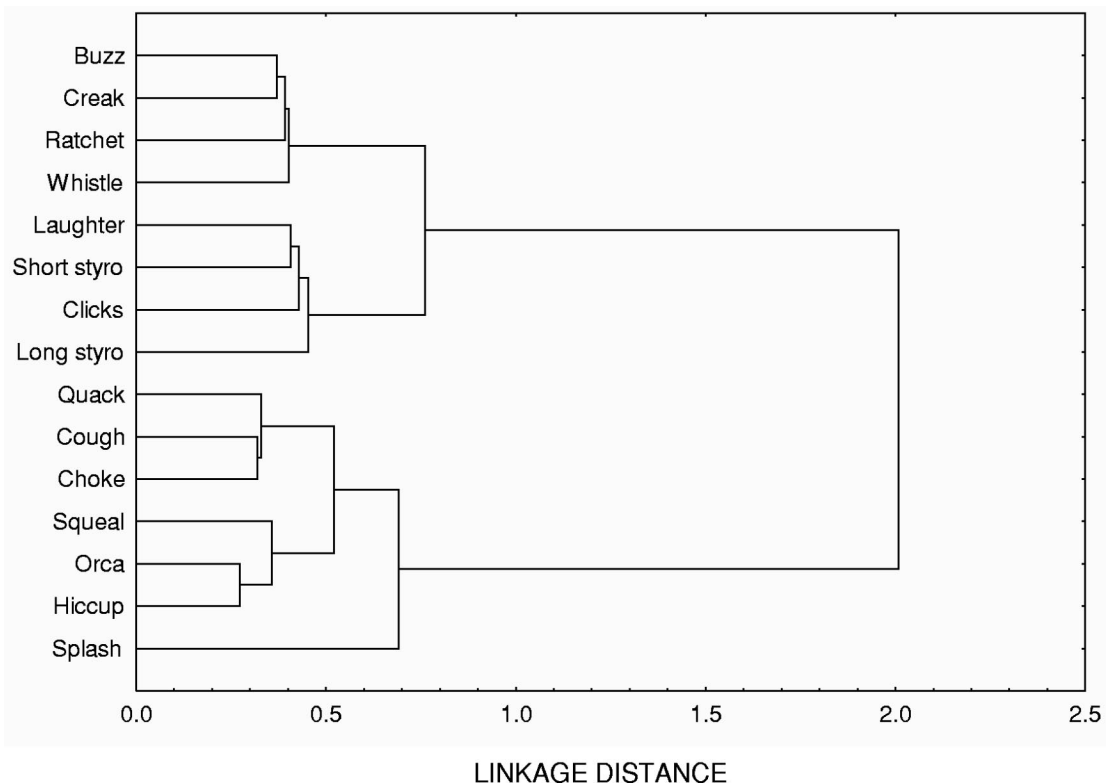


FIG. 3. Dendrogram derived from hierarchical cluster analysis (Ward's method). Based on dissimilarity matrix of SPCC scores.

burst pulses. Nevertheless, they appear in the nonrepetitive cluster. These two vocalizations form the strongest cluster in the dendrogram with a linkage distance of 0.27, perhaps reflecting their repetitive nature. *Orca* and *hiccup* also form a tight group with *squeal*. Despite being a tonal signal, *squeal* is aurally very similar to *orca* and this similarity is reflected in the tight clustering of these signals.

E. Principal component analysis of signal type

There was some correlation between the six parameters common to all 1959 vocalizations, and the use of PCA was therefore appropriate. The eigenvalues for the first three components were all greater than one, a value expected by chance alone for each variable. The resulting principal components accounted for 71% of the original variation, suggesting that the complexity of the data set can be reduced to three components with a 29% loss of information. This process formed separate clusters for *whistle*, *splash*, and *click train* (Fig. 4). The remaining signals form two general data clouds: the single bursts and those vocalizations with repeated elements.

Component 1 is most closely correlated with the variables “center time” and “peak time.” This component readily distinguished *splash* from the other vocalizations, and to a certain extent *ratchet*. *Splash* is a loud, broadband signal that seems to be produced nonvocally, typically through aerial activity and rapid movements of a dolphin's tailstock (personal observation). After a rapid onset, the energy produced during a *splash* degrades slowly, presumably due to surface echo and multipath effects. Thus the energy distribution in a typical *splash* is negatively skewed, with a

relatively low center time and peak time. Conversely, *ratchets* have high PC1 scores as they tend to increase in amplitude between successive clicks and are thus positively skewed. Component 3 readily distinguishes *whistles* from the other signal types and is strongly correlated with “peak frequency” (positively) and “interpeak interval” (negatively). Interpeak interval is a measure of the wavelength of a signal (i.e., the time offset between successive pressure peaks) and so, as wavelength decreases, frequency increases. *Whistles* have correspondingly high component 3 scores. Component 2 is strongly correlated with total signal duration and PC2 scores for the sequenced sounds (*laughter*, *short styro*, and *long styro*) are all very high. Many of these vocalizations are

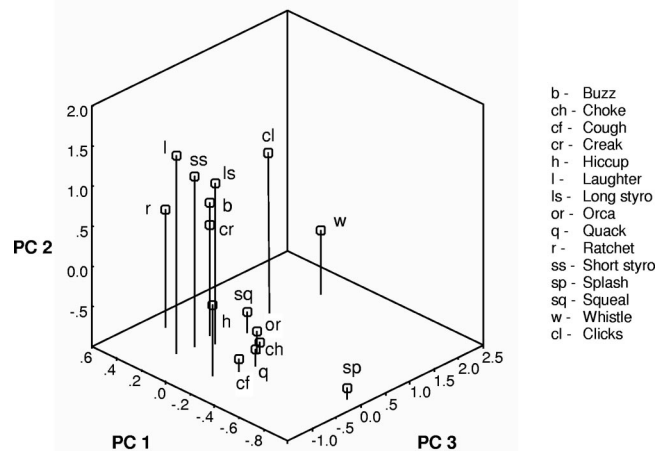


FIG. 4. Mean scores for first three principal components for each proposed signal type.

TABLE IV. Stepwise formation of an increasingly condensed repertoire. The final four categories represent fundamentally different structural classes (“tonal,” “single pulse,” “click bursts,” and “repeat bursts”).

Stepwise no. of signal types	% correct		Merging process	
	Observed	Expected	Candidate	Merged with
15	37.6	6.7		
14	38.6	7.1	long styro	short styro
13	40.9	7.7	cough	quack
12	43.9	8.3	laughter	long styro+short styro
11	45.3	9.1	choke	cough+quack
10	48.5	10.0	buzz	creak
9	47.5	11.1	orca	buzz+creak
8	55.3	12.5	ratchet	orca+buzz+creak
7	55.5	14.3	hiccup	laughter+long styro+short styro
6	62.6	16.7	splash	choke+cough+quack
5	65.7	20.0	clicks	ratchet+orca+buzz+creak
4	73.8	25.0	squeal	whistle

particularly long—for example, one bout of *laughter* used in this analysis was audible for 38 s and consisted of 130 bursts. Although not included in this analysis, some bouts may last longer than 2 min. As a result, these sequenced sounds form a subcluster along the PC2 axis. *Whistles* are separated from other vocalizations along this axis as they also tend to be long signals (the maximum duration in this data set was 3.3 s).

The first three principal component scores for all 1959 cases were used in discriminant function analysis (DFA). PCA scores were used in lieu of raw data to ensure variables were not correlated. The mean values for the three components differed significantly between signal types (ANOVA; all $F > 15$, all $p < 0.01$). It may therefore be assumed that all three components contribute to the discriminant model. To test the efficacy of the classification scheme, all 1959 cases were cross-validated. As there were 15 proposed signal types, a correct classification rate of approximately 7% would be expected by chance alone. The DFA process classified 37% of cases to the correct signal type. *Whistles* were most reliably allocated to the correct category (94% correct). *Splash* was also relatively well discriminated (77%), while *hiccup*, *laughter*, *orca*, *ratchet*, and *squeal* were also considered fairly distinct (all over 40% correct). The remaining vocalizations were poorly discriminated, with values ranging from 10% to 30% correct. *Long styro* had no correct classifications, and tended to be mistaken for the other sequenced vocalizations *laughter* and *ratchet*.

F. Condensation of signal repertoire

The results of the various multivariate techniques suggest that the variation of the measured parameters did not accurately capture the proposed repertoire. There was a great deal of overlap between some of the subjectively defined signal types and a clear need to reduce redundancy in the proposed repertoire. The results from the various multivariate techniques were used to construct a hierarchy of vocalizations that could be condensed with others in a stepwise process to gradually reduce the 15 initial groups to the four structural classes of “tonal,” “single pulse,” “click bursts,” and “repeat bursts” (Table IV). For each step the ratio between observed and expected classification accuracy was calculated.

A plot of these ratios against category number revealed a gradual increase followed by a plateau between 12 and 15 categories (Fig. 5). A plateau of this nature represents repertoire redundancy, as an increase in category number gives little net increase in classification success. Thus the 12 category repertoire appears to have maximal classification success and minimal redundancy.

The first step in the 12-signal repertoire was merging the two *styro* categories. The two categories had proved very hard to distinguish statistically. Indeed, during DFA no cases of *long styro* were correctly assigned. The second step was to join *cough* and *quack* into a new category defined as the

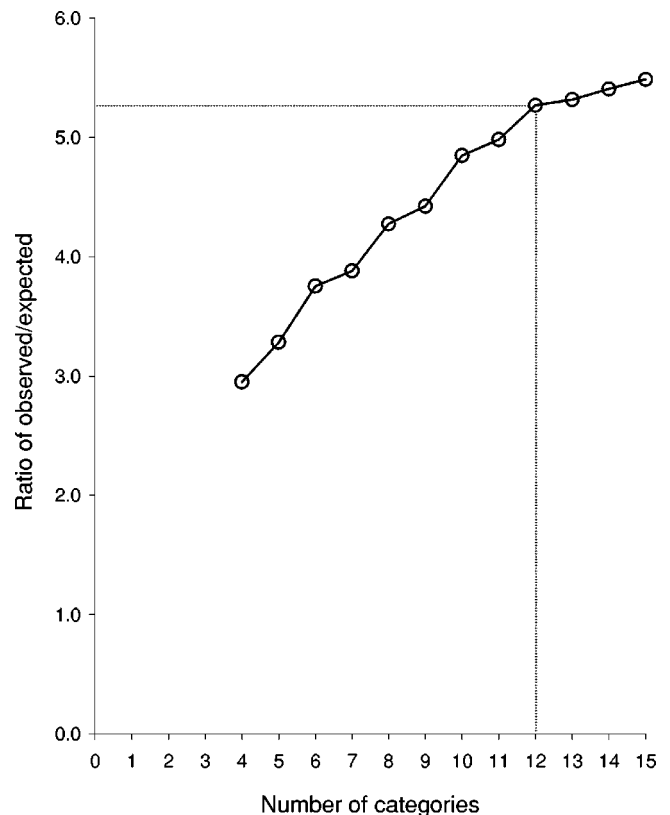


FIG. 5. Increase in correct classification (ratio of observed to expected values) with increasing number of signal types. Dotted line indicates the ratio of a 12-signal repertoire.

short bursts. These two vocalizations generally have very similar parameters and often clustered together during testing. The last step was to merge *laughter* with the *styros* into the category *sequenced signals*. Although these vocalizations were structurally different (i.e., single, double, and triple repeats), they were often confused during multivariate ordination. The 12-signal repertoire had a correct classification rate of 44% compared to 8.3% expected by chance (giving an observed to expected ratio of 5.3).

IV. DISCUSSION

The acoustic repertoire of the Fiordland dolphins is extremely diverse. Many of the vocalizations reported are structurally similar to signals reported elsewhere (Table I). *Whistles* and *click trains*, for example, have been reported from all acoustic studies of the *Tursiops* genus. Conversely, some vocalizations reported in other areas are not heard in the fiords, such as the low-frequency narrow-band calls recorded off eastern Australia (Schultz *et al.*, 1995). It would appear there is a geographical element to vocal repertoire that may in turn be based on differences in social structure and habitat type.

Broadband studies of cetacean repertoire are rare, particularly for free-ranging populations. High-frequency recordings of bottlenose dolphins in this study suggest that extreme broadband energy is restricted to click-based vocalizations and whistle harmonics. The click-based vocalizations may be important for echolocation, while the lower frequency vocalizations may have more communicative significance. The nature of sound is such that decreasing wavelength allows finer resolution of echoes reflected from small targets (Au, 1993). Hence echolocation clicks typically encompass higher frequencies, presumably for detecting and tracking small prey items. Conversely, lower frequency signals attenuate less energy and are more suited for long-range transmission of information during interspecific communication. In general, high frequencies are less ubiquitous for the Fiordland dolphins than for other species for which ultrasonic signals have been quantitatively described [e.g., Hector's dolphin (Dawson and Thorpe, 1990); harbour porpoise (Verboom and Kastelein, 1997); spinner and spotted dolphins (Lammers *et al.*, 2003)]. This difference may relate to relative body size. The four species mentioned are all small odontocetes, typically less than 2 m in length. The bottlenose dolphins in this study are significantly larger, with asymptotic body length conservatively predicted at 3.2 m (Schneider, 1999).

Categorization of vocalizations within a species' repertoire is essential in order to facilitate insight into functionality, social relevance, and geographical variation. Historically, dolphin vocalizations have been categorized as either tonal or pulsed (e.g., Herman and Tavolga, 1980). Some studies have further divided the pulsed category into "click trains" and "burst pulses" (Lilly and Miller, 1961; Caldwell and Caldwell, 1967; Ridgway, 1983). Results from this study are in broad agreement, with "tonal signals," "single bursts," and "click bursts" considered structurally different. However, a fourth division, "repeat bursts," is also suggested to further discriminate the pulsed vocalizations. This division is

based on the relative distinctiveness of these signals during multivariate testing combined with the growing evidence of sequenced signals from other bottlenose populations. It appears that the acoustic repertoire of bottlenose dolphins lies on a continuum with *clicks* at one extreme and *whistles* at the other. As interburst interval increases, *click trains* become *creaks*, *creaks* become *buzzes*, and so on. As consecutive pulses overlap to a greater and greater extent, the waveform becomes sinusoidal and eventually results in a *whistle*. Within the four structural classes further subdivision is possible. The initial repertoire of 15 signals was subsequently condensed to 12 based on various classification techniques. Throughout these ordinations, *whistles* were found to be uniformly different from all other signals, even the related *squeals*. Surprisingly, *splash* was also unique despite having similar aural properties to the *chokes*. Perhaps less surprising were the numerous differences between the *clicks* and all other vocalizations. The sequenced signals also tended to cluster together as might be expected. The other proposed signal types were less obviously demarcated. Condensation within the "single burst" and "repeat burst" classes provided a simplification of the repertoire with minimal change in accuracy.

Animal species tend to have small repertoires of signal units. A species' total repertoire for all sensory modalities is typically 40 to 45 units (Smith, 1969). Small repertoire size promotes the evolution of context-dependent communication. Within the cetaceans this is most obvious in those species limited to click-based vocalizations—the physeteridae (Gordon, 1987), the phocoenidae (Mohl and Andersen, 1973), and the genus *Cephalorhynchus* (Dawson, 1991). A similar scenario is played out in the repertoire of bottlenose dolphins. The click-based vocalizations have historically been associated with echolocation for navigation, orientation, and prey detection. Recent studies, however, suggest click-based vocalizations may also be socially important, such as the "pops" and "genital buzzes" produced by male dolphins during consortship [Connor and Smolker (1996) and Herzing (1996), respectively]. Thus the small repertoire of 12 signals proposed in this study may have expanded relevance in differing behavioural context. For example, a *buzz* produced in a foraging bout may represent the detection of prey during a rapid approach. The same vocalization produced in a social setting may be used for genital stimulation during sexual interactions.

The various classification techniques used in this study had different strengths, and thus often gave conflicting results. The SPCC process, for example, while providing a quantitative and repeatable comparison measure without subjective influence, has a number of weaknesses. Studies on the vocal output of songbirds suggest that as the background noise of recordings increase, between-group similarity also increases. This in turn leads to a reduction in SPCC accuracy (Terry *et al.* 2001). The SPCC technique also seems less reliable when comparing signals of different duration (Cortopassi and Bradbury, 2000). In general, SPCC techniques appear most useful when comparing similar signals, such as whistle contours. For the multivariate techniques, a drawback of typical PCA methods is that they are least squares

estimation techniques. Hence they are sensitive to outliers, which are common in realistic training sets (Kendall, 1980). In this study, the PCA process built upon quantitative measures taken automatically by the *Canary* software. As such, it was based upon subjectively selected parameters that were deemed important to human observers. However, problems of this kind are largely insurmountable. For example, the vocalizations in this study had wildly differing durations ranging from the most diminutive *splash* (0.02 s) to the most extravagant sequenced signal (120 s). This is the fundamental nature of these vocalizations and will always make classification difficult. The repertoire of vocalizations described in this paper was classified using a combination of human interpretation and multivariate ordination. However, the acoustic criteria important for signal representation in the dolphin auditory cortex remain unclear. Nevertheless, the 12-signal repertoire proposed in this study may provide the basis for subsequent analyses of geographical variation, functional importance, spatial and temporal patterns, and call sequencing.

ACKNOWLEDGMENTS

Heartfelt thanks are extended to all the volunteers who gave generously of their time to help in the field. The assistance and thoughtful suggestions of David Lusseau, Susan Lusseau, Dave Rundgren, and Steve Dawson were of particular importance. This work was made possible through the support of the New Zealand Whale and Dolphin Trust, New Zealand's Department of Conservation, the University of Otago and Real Journeys Ltd. Constructive comments were provided by two reviewers and greatly improved the original manuscript.

Au, W. W. L. (1993). *The Sonar of Dolphins* (Springer-Verlag, New York).
 Boisseau, O. J. (2005). "The Acoustic Behavior of Resident Bottlenose Dolphins in Fiordland, New Zealand." Ph.D. Thesis, Univ. of Otago.
 Bradbury, J. W., and Vehrencamp, S. L. (1998). *Principles of Animal Communication* (Sinauer Associates, Sunderland, MA).
 Brownlee, S. M., and Norris, K. S. (1994). "The acoustic domain," in *The Hawaiian Spinner Dolphin*, edited by K. S. Norris, B. Wursig, R. R. Wells, and M. Wursig (Univ. of California, Berkeley), pp. 161–185.
 Budde, C., and Klump, G. M. (2003). "Vocal repertoire of the black rhino *Diceros bicornis* and possibilities of individual identification," *Mammal Biol.* **68**(1), 42–47.
 Caldwell, M. C., and Caldwell, D. K. (1967). "Intraspecific transfer of information via the pulsed sound in captive odontocete cetaceans," in *Animal Sonar Systems: Biology and Bionics*, edited by R. G. Busnel (Laboratoire de Physiologie Acoustique, Jouy-en-Josas), pp. 879–936.
 Chabot, D. (1988). "A quantitative technique to compare and classify humpback whale (*Megaptera novaeangliae*) sounds," *Ethology* **77**(2), 89–102.
 Clark, C. W. (1982). "The acoustic repertoire of the southern right whale, a quantitative analysis," *Anim. Behav.* **30**, 1060–1071.
 Connor, R. C., and Smolker, R. A. (1996). "'Pop' goes the dolphin: A vocalization male bottlenose dolphins produce during consortships," *Behaviour* **133**, 643–662.
 Connor, R. C., Heithaus, M. R., Berggren, P., and Miksis, J. L. (2000). "'Kerplunking': Surface fluke-splashes during shallow-water bottom foraging by bottlenose dolphins," *Marine Mammal Sci.* **16**(3), 646–653.
 Corkeron, P. J., and Van Parijs, S. M. (2001). "Vocalizations of eastern Australian Risso's dolphins, *Grampus griseus*," *Can. J. Zool.* **79**(1), 160–164.
 Cortopassi, K. A., and Bradbury, J. W. (2000). "The comparison of harmonically rich sounds using spectrographic cross-correlation and principal coordinates analysis," *Bioacoustics* **11**, 89–127.

Dawson, S. M., and Thorpe, C. W. (1990). "A quantitative analysis of the sounds of Hector's dolphin," *Ethology* **86**, 131–145.
 Dawson, S. M. (1991). "Clicks and communication—the behavioural and social contexts of Hector's dolphin vocalisations," *Ethology* **88**(4), 265–276.
 dos Santos, M. E., Ferreira, A. J., and Harzen, S. (1995). "Rhythmic sound sequences emitted by aroused bottlenose dolphins in the Sado estuary, Portugal," in *Sensory Systems of Aquatic Mammals*, edited by R. A. Kastelein, J. A. Thomas, and P. E. Nachtigall (De Spil, Woerden), pp. 325–334.
 dos Santos, M. E., Caporin, G., Moreira, H. O., Ferreira, A. J., and Coelho, J. L. B. (1990). "Acoustic behavior in a local population of bottlenose dolphins," in *Sensory Abilities of Cetaceans: Laboratory and Field Evidence*, edited by J. A. Thomas and R. A. Kastelein (Plenum, New York), pp. 585–598.
 Finneran, J. J., Oliver, C. W., Schaefer, K. M., and Ridgway, S. H. (2000). "Source levels and estimated yellowfin tuna (*Thunnus albacares*) detection ranges for dolphin jaw pops, breaches, and tail slaps," *J. Acoust. Soc. Am.* **107**, 649–656.
 Fischer, J., and Hammerschmidt, K. (2002). "An overview of the Barbary macaque, *Macaca sylvanus*, vocal repertoire," *Folia Primatol.* **73**(1), 32–45.
 Ford, J. K. B. (1989). "Acoustic behaviour of resident killer whales (*Orcinus orca*) off Vancouver Island, British Columbia," *Can. J. Zool.* **67**, 727–745.
 Gish, S. L. (1979). "A quantitative description of two-way acoustic communication between captive Atlantic bottlenosed dolphins (*Tursiops truncatus Montagu*)," Ph.D. thesis, University of California at Santa Cruz.
 Goodson, A. D., Klinowska, M., and Morris, R. (1988). "Interpreting the acoustic pulse emissions of a wild bottlenose dolphin (*Tursiops truncatus*)," *Aquat. Mamm.* **14**(1), 7–12.
 Gordon, J. C. (1987). "The behaviour and ecology of sperm whales off Sri Lanka," Ph.D. thesis, University of Cambridge.
 Grafe, T. U., Steffen, J. O., and Stoll, C. (2000). "Vocal repertoire and effect of advertisement call intensity on calling behaviour in the West African tree frog *Leptopelis viridis*," *Amhibi reptilia* **21**(1), 13–23.
 Hastie, G. D., Wilson, B., and Thompson, P. M. (2003). "Fine-scale habitat selection by coastal bottlenose dolphins: application of a new land-based video-montage technique," *Can. J. Zool.* **81**(3), 469–478.
 Herman, L. M., and Tavolga, W. N. (1980). "The communication systems of cetaceans," in *Cetacean Behavior: Mechanisms & Functions*, edited by L. M. Hermans (Wiley, New York), pp. 149–210.
 Herzing, D. L. (1996). "Vocalizations and associated underwater behavior of free-ranging Atlantic spotted dolphins *Stenella frontalis* and bottlenose dolphins, *Tursiops truncatus*," *Aquat. Mamm.* **22**(2), 61–79.
 Herzing, D. L. (2000). "Acoustics and social behaviour of wild dolphins: Implications for a sound society," in *Hearing by Whales and Dolphins*, edited by W. W. L. Au, A. N. Popper, and R. R. Fay (Springer, New York), pp. 225–272.
 Jacobs, M., Nowacek, D. P., Gerhart, D. J., Cannon, G., Nowicki, S., and Forward, R. B. (1993). "Seasonal changes in vocalizations during behavior of the Atlantic bottlenose dolphin," *Estuaries* **16**(2), 241–246.
 Janik, V. M. (2000a). "Whistle matching in wild bottlenose dolphins (*Tursiops truncatus*)," *Science* **289**, 1355–1357.
 Janik, V. M. (2000b). "Food-related bray calls in wild bottlenose dolphins (*Tursiops truncatus*)," *Proc. R. Soc. London, Ser. B* **267**(1446), 923–927.
 Jones, G. J., and Sayigh, L. S. (2002). "Geographic variation in rates of vocal production of free-ranging bottlenose dolphins," *Marine Mammal Sci.* **18**(2), 374–393.
 Kendall, M. G. (1980). *Multivariate Analysis* (Macmillan, New York).
 Kroodsma, D. E., and Henri, O. (1982). *Acoustic Communication In Birds: Sounds Production, Perception And Design Features Of Sounds* (Academic Press, London).
 Lammers, M. O., Au, W. W. L., and Herzing, D. L. (2003). "The broadband social acoustic signaling behavior of spinner and spotted dolphins," *J. Acoust. Soc. Am.* **114**, 1629–1639.
 Lilly, J. C., and Miller, A. M. (1961). "Sounds emitted by the bottlenose dolphin," *Science* **133**, 1689–1693.
 McCowan, B., and Reiss, D. (1995). "Maternal aggressive contact vocalizations in captive bottlenose dolphins (*Tursiops truncatus*): wide-band, low-frequency signals during mother/aunt-infant interactions," *Zoo Biology* **14**(4), 293–309.
 Mohl, B., and Andersen, S. (1973). "Echolocation: high-frequency compo-

- ment in the click of the harbour porpoise (*Phocoena phocoena*)," J. Acoust. Soc. Am. **54**, 368–372.
- Monteiro, E. L. A., and Monteiro, K. (2001). "Low-frequency sounds emitted by *Sotalia fluviatilis guianensis* (Cetacea: Delphinidae) in an estuarine region in southeastern Brazil," Can. J. Zool. **79**(1), 59–66.
- Murray, S. O., Mercado, E., and Roitblat, H. L. (1998a). "Characterizing the graded structure of false killer whale (*Pseudorca crassidens*) vocalizations," J. Acoust. Soc. Am. **104**, 1679–1688.
- Murray, S. O., Mercado, E., and Roitblat, H. L. (1998b). "The neural network classification of false killer whale (*Pseudorca crassidens*) vocalizations," J. Acoust. Soc. Am. **104**, 3626–3633.
- Rasmussen, M. H., and Miller, L. A. (2002). "Whistles and clicks from white-beaked dolphins, *Lagenorhynchus albirostris* recorded in Faxaflói Bay," Aquatic Mammals **28**, 78–89.
- Recchia, C. A. (1994). "Social Behaviour of Captive Belugas, *Delphinapterus leucas*," Ph.D. thesis, Massachusetts Institute of Technology.
- Reiss, D. (1988). "Observations on the development of echolocation in young bottlenose dolphins," in Animal Sonar: Processes and performance, edited by P. E. Nachtigall and P. W. B. Moore (Plenum, New York), pp. 121–127.
- Richardson, W. J., Thomson, D. H., Greene, C. R., and Malme, C. I. (1995). "Marine mammal sounds," in *Marine Mammals and Noise*, edited by W. J. Richardson, D. H. Thomson, C. R. Greene, and C. I. Malme (Academic, London), pp. 159–204.
- Ridgway, S. H. (1983). "Dolphin hearing and sound production in health and illness," in *Hearing and Other Senses: Presentations in Honor of E. G. Wever*, edited by R. R. Fay and G. Gourevitch (Amphora, Groton, CT), pp. 247–296.
- Sayigh, L. S., Tyack, P. L., Wells, R. S., Solow, A. R., Scott, M. D., and Irvine, A. B. (1999). "Individual recognition in wild bottlenose dolphins: a field test using playback experiments," Anim. Behav. **57**, 41–50.
- Schneider, K. (1999). "Behaviour and ecology of bottlenose dolphins in Doubtful Sound, Fiordland, New Zealand," Ph.D. thesis, Univ. of Otago.
- Schultz, K. W., Cato, D. H., Corkeron, P. J., and Bryden, M. M. (1995). "Low frequency narrow-band sounds produced by bottlenose dolphins," Marine Mammal Sci. **11**(4), 503–509.
- Sjare, B. L., and Smith, T. G. (1986). "The vocal repertoire of white whales, *Delphinapterus leucas*, summering in Cunningham Inlet, Northwest Territories," Can. J. Zool. **64**, 407–415.
- Smith, W. J. (1969). "Messages of vertebrate communication," Science **165**, 145–150.
- Smolker, R., and Pepper, J. W. (1999). "Whistle convergence among allied male bottlenose dolphins (Delphinidae, *Tursiops* sp.)," Ethology **105**(7), 595–617.
- Stevens, J. P. (2001). *Applied Multivariate Statistics for the Social Sciences (Fourth Edition)* (Erlbaum, Hillsdale, NJ).
- Terry, A. M. R., McGregor, P. K., and Peake, T. M. (2001). "A comparison of some techniques used to assess vocal individuality," Bioacoustics **11**, 169–188.
- Van Parijs, S. M., and Corkeron, P. J. (2001). "Vocalizations and behaviour of Pacific humpback dolphins *Sousa chinensis*," Ethology **107**(8), 701–716.
- Verboom, W. C., and Kastelein, R. A. (1997). "Structure of harbour porpoise (*Phocoena phocoena*) click train signals," in *The Biology of the Harbour Porpoise*, edited by A. J. Read, P. R. Wiepkema, and P. E. Nachtigall (De Spil, Woerden), pp. 343–362.
- Volodina, E. V. (2000). "Vocal repertoire of the cheetah *Acinonyx jubatus* (Carnivora, Felidae) in captivity: Sound structure and their potential for estimating the state of adult animals," Zool Zh. **79**(7), 833–843.
- Ward, J. H. (1963). "Hierarchical grouping to optimize objective function," J. Am. Stat. Assoc. **58**, 236–244.
- Watkins, W. A., Schevill, W. E., and Best, P. B. (1977). "Underwater sounds of *Cephalorhynchus heavisidii* (Mammalia: Cetacea)," J. Mammal. **58**, 316–318.
- Wood, F. G. (1953). "Underwater sound production and concurrent behavior of captured porpoises, *Tursiops truncatus* and *Stenella plagiodon*," Bull. Mar. Sci. Gulf Caribb. **3**, 120–133.
- Zar, J. H. (1998). *Biostatistical Analysis (4th Edition)* (Prentice Hall, Englewood Cliffs, NJ).

Information theory analysis of patterns of modulation in the advertisement call of the male bullfrog, *Rana catesbeiana*^{a)}

Dianne N. Suggs and Andrea Megela Simmons^{b)}

Department of Psychology, Brown University, Providence, Rhode Island 02912

(Received 7 October 2004; revised 10 January 2005; accepted 11 January 2005)

Male bullfrogs often amplitude modulate the envelopes of the individual notes (croaks) in their multinote advertisement calls. These amplitude modulations change the envelope of the note from smooth and unmodulated to one with varying numbers of modulations. A Markov analysis shows the pattern of change in the envelope to be highly ordered, but not completely so (semi-Markovian). Three simple rules govern the presence or absence of modulations in individual notes. These rules are (1) all calls begin with an unmodulated note; (2) the first note to be modulated will contain only one modulation; and (3) when a change in modulation occurs from one note to the next, it does so with an increase or a decrease of one modulation only. The addition of modulations is correlated with an increase in note duration. Physiologically, the presence of modulations might increase the precision of temporal coding of note periodicities in the central auditory system. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1863693]

PACS numbers: 43.80.Ka [WA]

Pages: 2330–2337

I. INTRODUCTION

Animals use acoustic signals of different form and complexity to communicate such messages as species identification, individual identification, advertisement of social status, advertisement of biological fitness, sexual motivation, and aggressive motivation. Some species have evolved large, complex repertoires to communicate these different messages, while others use a more limited repertoire. A fundamental question in the study of acoustic communication concerns how animals are able to convey multiple, biologically relevant messages with a limited number of signals. The acoustic behavior of anuran amphibians (frogs and toads) has been extensively studied as a model of a communication system consisting of a relatively small repertoire (Wells, 1977; Gerhardt and Huber, 2002). In some species, separate notes in a call communicate distinct messages to different receivers (Narins and Capranica, 1978), while in other species, the same notes may send simultaneous, different messages to different receivers (Capranica, 1965). The incorporation of different messages within a limited repertoire may be achieved by varying the complexity as well as the degree of stereotypy of the signal (Gerhardt and Huber, 2002).

The male bullfrog (*Rana catesbeiana*) emits a harmonically complex advertisement call that is important for regulating both male–male territorial and male–female attractive interactions (Emlen, 1968; Wiewandt, 1969). The spectral and temporal structure of this vocalization has been described in several publications (Capranica, 1965; Wiewandt, 1969; Bee and Gerhardt, 2001; Bee, 2004; Simmons, 2004). The advertisement call consists of a series of 1–12 individual notes (croaks). The one-note call has been categorized as having a territorial or aggressive function (Capranica, 1965),

but the communicative function of the variation in number of notes beyond one has not been definitively established. Each note consists of a number of harmonically related frequencies, extending from about 200 to about 2000 Hz, with a missing or attenuated fundamental frequency around 100 Hz. There is considerable interindividual variability in acoustic properties of individual notes (Bee and Gerhardt, 2001; Bee, 2004). Males may be able to use these differences to discriminate neighbors from strangers, and near neighbors from far neighbors (Davis, 1987; Boatright-Horowitz *et al.*, 2000; Bee and Gerhardt, 2001). Which features female bullfrogs use to choose particular males in a chorus has not been established.

The purpose of this study is to describe and analyze a feature of the male bullfrog's advertisement call that has not as yet been addressed in the literature. Whereas the envelope, or overall shape, of the individual notes has typically been represented as smooth and unmodulated (Capranica, 1965; Bee and Gerhardt, 2001; Bee, 2004; Simmons, 2004), this is frequently not the case. Rather, males often add extra modulations to the envelopes of the individual notes in their calls. The effect of these extra modulations is to change the shape of the envelope so that visually it contains partial gaps, or repetitive valleys. Human listeners perceive these gaps as slight hesitations or pauses (“stutters”). Using analyses based on information theory (Shannon and Weaver, 1949), we characterize the pattern of change in the envelope of individual notes in the advertisement call. Our data show that modulations are added to individual notes in a highly ordered fashion. The robustness of the pattern suggests that these modulations may serve some communicative function.

II. MATERIALS AND METHODS

A. Field sites

Recordings of the calling activity of male bullfrogs from 1994 ($N=10$) and 2003 ($N=10$) were used to analyze the

^{a)}Portions of this work were presented at the 75th Anniversary Meeting of the Acoustical Society of America, May 2004, and the 2004 Annual Meeting of the Animal Behavior Society.

^{b)}Electronic mail: andrea_simmons@brown.edu

presence of modulations in individual notes in their advertisement calls. In July 1994 spontaneous chorusing was recorded at a natural bullfrog breeding pond (80×35 m) located in eastern Rhode Island (41°44'34"N,71°20'14"W); analysis of the vocal interactions between males in this chorus has been published previously (Boatright-Horowitz *et al.*, 2000). Air and water temperatures were measured to an accuracy of 0.5 °C during each recording session and ranged between 21 and 25 °C. Chorus density varied from eight to ten vocalizing males each night.

In June and July 2003, spontaneous chorusing was recorded (as part of a larger experiment, results of which will be presented separately) at three breeding sites in Rhode Island and Massachusetts. The location of the 1994 recordings served as site 1 in the 2003 recording sessions. Chorus density varied between six and eight vocalizing males. Site 2, located in central Massachusetts (42°25'47"N, 71°49'35"W), is a permanent pond approximately 100 m in length and 30 m wide. Chorus density ranged from six to ten vocalizing males. Site 3, located in northwestern Rhode Island (41°45'00"N,71°45'00"W), is a permanent pond approximately 90 m long and 25 m wide. Chorus density was lower at this site, ranging from three to five calling males per night. Air and water temperatures of the three sites were measured to an accuracy of 0.5 °C and ranged between 17 and 26 °C.

B. Data collection

During 1994, spontaneous chorus activity was recorded on 3 nights between the hours of 22.00 and 00.15 h. Male advertisement calls were recorded using a battery-powered Marantz cassette recorder (model PMD430: frequency response 0.05–14 kHz, ±3 dB) and a Bionic Ear microphone (frequency response 0.1–14 kHz, ±3 dB) positioned on a Styrofoam platform placed in the water 1 m from an actively calling male. In 2003, chorus activity was recorded on 10 nights between the hours of 21.00 and 01.00 h using a Midnight Blue MB2000L (frequency response 0.06–15 kHz, ±3 dB) or Sennheiser ME66 (frequency response 0.05–20 kHz, ±3 dB) microphone, positioned on a Styrofoam platform about 1 m from individual frogs, and recorded onto the Marantz cassette recorder. Recording sessions lasted between 30 min and 3 h depending on the activity of the chorus. Individual males within each chorus were identified through field notes and visual and auditory inspection of each individual's calls. Such identification is possible due to the small amount of variation in fundamental and dominant frequencies within an individual's calls, whereas significant variability can be detected between the calls of individual frogs (Bee and Gerhardt, 2001). In addition, calling males maintained relatively fixed territorial positions during a single night's chorus and were easily identified by spatial location.

C. Data analysis

Spontaneous vocalizations were converted to 44.1-kHz, 16-bit WAV-format sound files using COOLEEDIT 2000. Figure 1 is an oscillogram of an advertisement call recorded during the summer of 1994. This particular call consist of six indi-

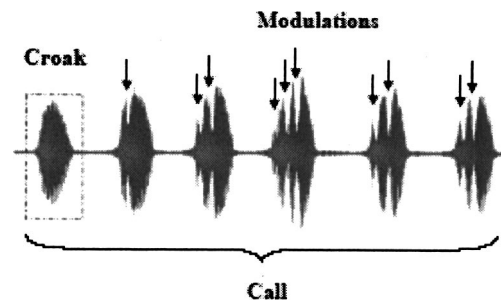


FIG. 1. Oscillogram of a male bullfrog's call recorded at site 1 in 1994. This particular call is made up of six individual notes or croaks. Arrows point to modulations or gaps seen in individual notes. The first note contains no modulations; notes two through six contain 1, 2, 3, 2, and 2 modulations, respectively. Time-domain waveform was exported from COOLEEDIT 2000.

vidual notes. The first note has a smooth, unmodulated envelope, while the envelopes of the other five notes contain modulations. For each call, the numbers of notes and the number of modulations within each note were determined in COOLEEDIT by visual and auditory inspection. The frequency of modulation, the fundamental frequency, and the frequency spectrum of each note were analyzed by a custom-written MATLAB v 6.5 (MathWorks, Natick, MA) script. Duration of individual notes was measured using COOLEEDIT (1-ms resolution) and analyzed using SPSS v. 11.5 statistical packages.

To analyze any pattern in the incorporation of modulations within notes, all advertisement calls with no modulations in any notes were eliminated. Calls whose individual notes could not be completely analyzed due to overlap by calls of neighboring males or interference by nonbiological sources of noise were eliminated as well. The remaining calls were analyzed using UNCERT, a program that examines serial dependencies in sequences of events using Markov chain analysis (Hailman and Hailman, 1993). The program calculates frequencies and probabilities of occurrence of a series of events, the frequencies and probabilities of transition from one event to another, and the uncertainties associated with the events and their transitions. These analyses can be used to evaluate the organizational complexity of, for example, a signal repertoire (McGowan *et al.*, 1999). In our data set, the sequence of events is the number of modulations within consecutive notes in a call. The hypothesis tested is that the number of modulations changes in a systematic way from note to note.

The diversity in the series of events is represented by an uncertainty value, a logarithmic measure (Hailman and Hailman, 1993) based on communication theory (Shannon and Weaver, 1949). To find the maximum amount of uncertainty in a data set, the following formula is applied: $U_m = \log_2 K$, where U_m measures the diversity of events within a sample, \log_2 is the logarithm to the base 2, and K denotes the number of different kinds of events involved. Consider, for example, an animal's repertoire consisting of just one kind of event. The maximum uncertainty would be calculated as $U_m = \log_2(1) = 0$. This makes intuitive sense, because with only one event in a repertoire, there would be absolutely no uncertainty as to what kind of event would follow. Zero-order uncertainty (U_0) is a measure of the probabilities of each of

TABLE I. Bullfrog advertisement calls collected during 1994 and 2003 by date and site. The number of calls, number of calls with modulations (mods), and number of one-note calls are indicated for each recording session. More than 80% of the modulated calls conform to the two patterns identified.

Date	Site no.	Total calls	Calls with mods	% Conformation to pattern		One-note calls	
				“begins with 1 mod”	“+ or - 1 mod”	Total	% with mods
7/02/94	1	406	155	87.1	99.9	196	0
7/02/94	1	307	116	92.2	98.3	165	0
7/03/94	1	299	137	92.0	99.3	123	0
7/08/94	1	290	171	81.9	91.2	165	0
6/24/03	1	249	67	94.0	100	28	0
6/25/03	1	505	245	89.4	98.0	71	0
6/26/03	1	495	228	92.6	98.7	32	0
6/27/03	1	435	180	85.6	94.4	105	0
6/28/03	1	280	102	92.2	96.1	92	0
6/29/03	2	321	232	96.1	97.4	15	0
6/30/03	2	335	231	99.1	99.6	30	0
7/06/03	1	562	230	89.6	98.3	127	0
7/07/03	1	366	142	86.6	97.2	96	0
7/08/03	2	184	81	93.8	98.8	36	0
7/10/03	3	116	58	100	96.6	30	0
7/13/03	1	254	89	85.4	93.3	84	0
7/14/03	3	163	72	97.2	100	60	0

the kinds of events and thus the complexity of the repertoire. Within each data set of events, each individual event (i) has its own frequency of occurrence (F_i). The total number of events of all kinds is the sum of the total frequencies (N). This means that the probability that event i (or P_i) will occur equals the frequency of i (or F_i) divided by the total frequencies of all events (N). If all events are equally probable, $U_0 = U_m$. First-order uncertainty (U_1) measures the dependencies between first and second events, and second-order uncertainties (U_2) measure the transitional probabilities between event pairs and third events in sequences (Hailman and Hailman, 1993). Although higher measures of uncertainties exist, our data do not exhibit the complexity that would warrant application of them.

III. RESULTS

A. General patterns

A total of 5567 bullfrog calls was collected from 17 recording sessions. Of these calls, 2599 consisted of single- and multinote calls that contained no modulations. Within the remaining 2968 calls, 432 could not be completely analyzed due to interference or call overlap and were eliminated from further analysis. This left a total of 2536 calls made up of 13 081 individual notes that were further analyzed for any pattern in the presence of modulations. Results of analysis of these calls listed according to date and sampling site are shown in Table I.

All single-note calls ($N = 1455$) had unmodulated envelopes. On any recording night, between 27%–72% of all calls (mean 30.4%) contained at least one note with modulations. Between 81.9%–100% of calls that contained any notes with modulations began with one unmodulated note. The mean percent of calls conforming to this “addition of modulations begins with 1 modulation” rule was 91.46%

$\pm 5.1\%$. When the number of modulations increased or decreased between successive notes within a call, it did so by just one modulation in 91.2%–100% of the calls. A mean of $97.5\% \pm 2.5\%$ of calls followed the “modulation increases or decreases by one” rule.

The number of modulations overall varied from 1 to 6 in separate calls, with modulation frequencies all below 12 Hz. One animal produced 5 notes with 7, 8, or 9 modulations, but this occurred so rarely that these notes were excluded from analyses with UNCERT.

B. Uncertainty analyses

The data set was categorized into seven different kinds of events based on the number of modulations (0 to 6) observed in call notes. Based upon the analysis performed by UNCERT, Table II reveals the frequencies and probabilities of the occurrences of each of the seven kinds of events. Maximum uncertainty (U_m) as computed on seven events is 2.81 bits/event and zero-order uncertainty (U_0) is 2.03 bits/event. A bit is a contraction of computer language’s “binary digit” and is considered the basic unit of information. What this means is that the number of bits is the power to which the number 2 must be raised in order to give the number of

TABLE II. Zero-order analysis. Frequencies and probabilities of occurrence of seven kinds of events as calculated from UNCERT.

Event	Frequency	Probability
0	4411	0.3372
1	3613	0.2762
2	3401	0.2600
3	1338	0.1023
4	285	0.0218
5	29	0.0022
6	4	0.0003

TABLE III. First-order analysis. Frequencies (and probabilities) of transition from preceding to following events

Preceding event	Following event						
	0	1	2	3	4	5	6
0	1791 (0.4168)	2274 (0.5294)	220 (0.0512)	9 (0.0021)	2 (0.0005)	0 (0.00)	0 (0.00)
1	100 (0.0378)	1033 (0.3899)	1482 (0.5590)	32 (0.0121)	1 (0.0004)	0 (0.00)	0 (0.00)
2	10 (0.0043)	290 (0.1234)	1419 (0.6038)	624 (0.2655)	7 (0.0030)	0 (0.00)	0 (0.00)
3	5 (0.0050)	16 (0.0160)	274 (0.2740)	577 (0.5770)	128 (0.1280)	0 (0.00)	0 (0.00)
4	0 (0.00)	0 (0.00)	5 (0.0200)	95 (0.3800)	131 (0.5240)	19 (0.0760)	0 (0.00)
5	0 (0.00)	0 (0.00)	1 (0.0357)	1 (0.0357)	16 (0.5714)	8 (0.2857)	2 (0.0714)
6	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	2 (0.5000)	2 (0.5000)

equiprobable messages in a signal or a system (Lewis and Gower, 1980). Table III shows a matrix depicting the frequencies of the events as they transition either to the following event or as a repetition of themselves. To obtain a Markovian chain analysis, the frequencies in Table III were converted to probabilities to give a clearer picture of the pattern of insertion of modulations. Based upon those probabilities, U_1 is calculated as 1.31 bits/events, and U_2 as 1.27 bits/event.

Individual graphs for each event are shown in Fig. 2. For each of the graphs, the initial state (S) is identified by the number in the upper-right corner. The probability of the next event is illustrated for each of the seven possibilities (0 through 6). For example, the event state S-0 is followed by a repetition of itself with a probability of 0.42, by S-1 with a probability of 0.53, and by S-2 with a probability of 0.05. Event state S-1 is followed by S-0 with a probability of 0.04, by a repetition of itself by 0.39, by S-2 with a probability of 0.56, and by S-3 with a probability of 0.01. Inspection of the individual graphs shows the increased probabilities of an event being followed by itself, by the preceding event, or by the following event. This pattern is precisely predicted by the “number of modulations increases or decreases by one” rule.

This pattern of transition can be graphically represented through the use of kinematic graphing. Kinematic graphs represent the temporal or sequential ordering of events. Because a system can only be in one state at a time, each state must exclude all other states. In addition, the states must be discrete and easily discriminated from one another (Sustare, 1978). The present data follow these rules perfectly. A male bullfrog can produce only one event state (a single note) at one time, and each event state within its call is discernible from all others (zero through 6 modulations in individual notes). In our kinematic diagram (Fig. 3), circles are used to represent individual events. Events are categorized and labeled according to the number of modulations contained within a note, beginning with S-0 (unmodulated), S-1 (one modulation), S-2 (two modulations), and so forth up to and including S-6 (six modulations). Each time one state ends and is followed by another state, a transition is said to occur.

When a transition occurs, it is depicted by an arrow between the two state events. The direction of the arrow shows the temporal sequence of the events—always pointing away from the initial state and towards the following state. Recurrent arrows, which point from one state back to itself, represent a temporal repetition of that state. When all arrows are

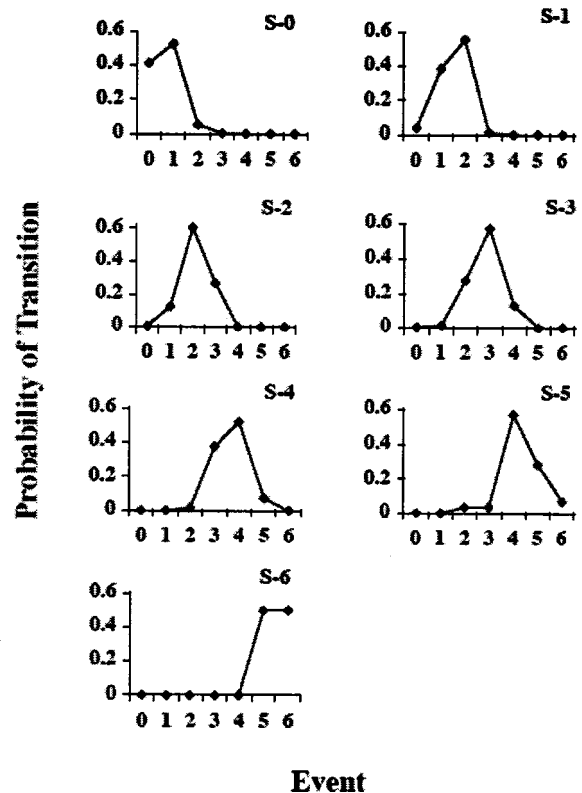


FIG. 2. Probabilities of transition from each of the seven kinds of events (S-0 to S-6, upper right corner in each graph) to each of the events (0 through 6 on x axis). For example, a note containing zero modulations (upper-left graph) is followed by another note with no modulations with a probability of 0.42, by a note with one modulation with a probability of 0.53, by a note with two modulations with a probability of 0.002, by a note with 4 modulations with a probability of 0.0005, and by a note with 5 or 6 modulations with a probability of 0.

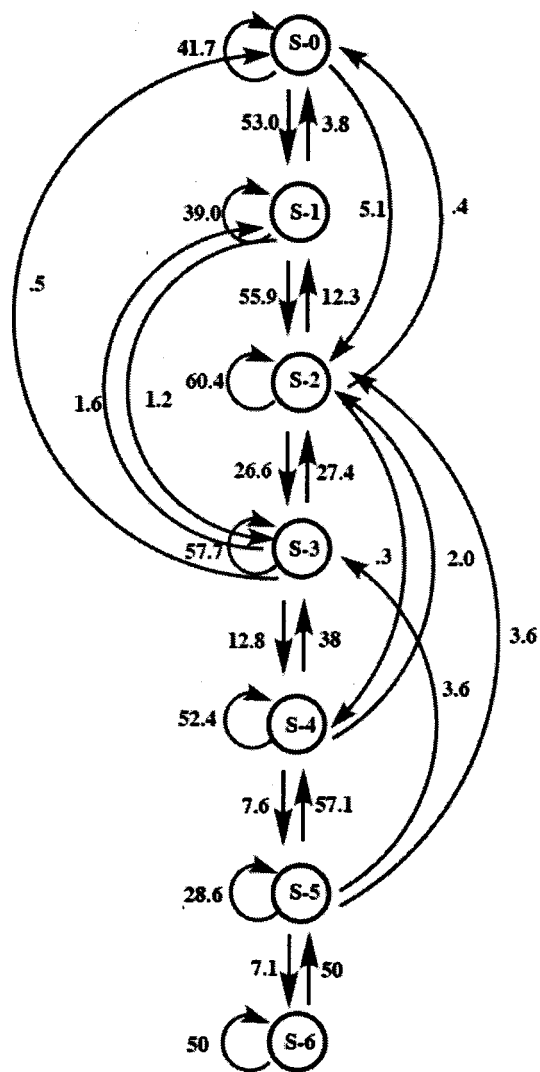


FIG. 3. Kinematic diagram of first-order Markov chain analysis. Events are represented by circles (S-0 through S-6, depending upon number of modulations contained within the note). Transitions between events are depicted by arrows pointing away from an initial event and towards a following event. Recurrent arrows represent repetitions of initial events. Probabilities of transitions are found next to arrows.

included in the kinematic graph, a complete picture of the behavior of the system results. Beginning with event S-0, 41.7% of the total S-0 events are recurrent (repeating); 52.9% transition to event S-1; 5.1% transition to event S-2; and very small fractions transition to events S-3 and S-4 (0.021% and 0.005%, respectively). Due to the relatively small amounts of information these transitions add to the system, coupled with the attempt to maintain clarity in the diagram, they are not represented by arrows in the kinematic graph. Moving through the graph, the same observations can be made for each of the event states (S-1 through S-6), with the frequencies of transitions emanating from each of the event states totaling 100%.

C. Correlation with other acoustic parameters

The addition of modulations to the envelope of individual notes is correlated with an increase in note duration.

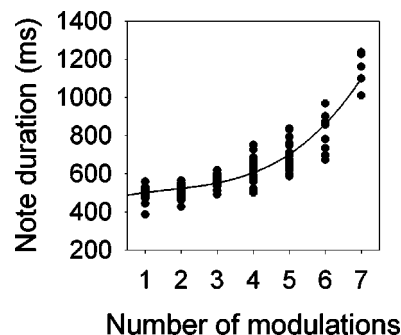


FIG. 4. Relationship between note duration and number of modulations per note in the spontaneous vocalizations of one male recorded at site 1 in 1994. The curved line is the best-fitting regression through the data.

This trend is shown in Fig. 4 for an individual male. Data are best fit by a cubic regression ($r^2=0.83, P<0.01$). Over all males, the relationship between note duration and number of modulations is statistically significant ($\rho=0.63, P<0.001$). In contrast, in advertisement calls with no modulated notes, there is no significant relationship ($\rho=0.12$) between note duration and its position within the call (that is, later notes do not have longer durations than earlier notes). The change in duration between successive notes thus is related to the insertion of modulations, not to note position.

Although the insertion of modulations into successive notes followed the patterns described for all males, there was individual variability in the extent to which males added modulations to their calls. Individual variability was determined for the ten males recorded in 1994, during sessions in which each male was recorded for periods of about 3 h. The percent of calls that included any modulations varied from 0% to 100% between individuals. There was no significant relationship ($r^2=-0.06$) between the percent of calls that included modulations and fundamental frequency of unmodulated notes in an individual's calls. From this analysis, we infer that the insertion of modulations may be related to chorus dynamics or to female choice rather than to male body size (which is correlated with fundamental frequency; Bee and Gerhardt, 2001).

D. Variation over chorus time

The numbers of calls containing modulations in any notes as a percent of total calls for all males from site 1 in 2003 are shown in Fig. 5(A). The percent of calls with modulations is low on the first recording day (which was the first day males were vocally active at that site), rises to a maximum of about 50% on the second day, and levels off around 40% until the end of the calling season at that site. Data from an individual male on one night, recorded at site 1 in 1994, are shown in Fig. 5(B). The number of notes with modulations (regardless of the number) given by this male varies over the entire chorusing bout, with a peak between calls 40 and 50. Notes with modulations are not restricted to either the beginning or the end of the calling bout, but occur throughout and seem to alternate with unmodulated calls.

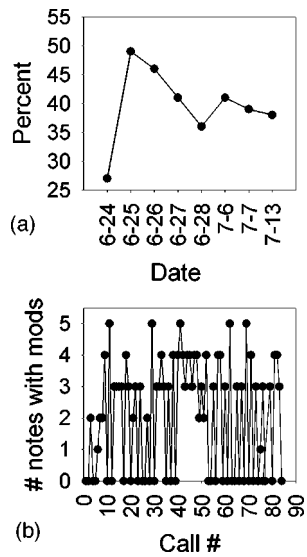


FIG. 5. (A) Percent of calls containing any modulations as a function of all calls in the spontaneous vocalizations of male bullfrogs recorded at site 1 in 2003. (B) Number of notes with modulations in the spontaneous calling of one male recorded at site 1 in 1994.

IV. DISCUSSION

Individual notes in the advertisement call of male bullfrogs often contain extra envelope modulations that visually appear as gaps within an individual note and which acoustically to human listeners resemble pauses or “stutters.” There appear to be three simple rules that govern the pattern of incorporation of these modulations within individual notes in an advertisement call. First, all calls begin with an unmodulated note. Second, the first note to contain a modulation contains only one. Third, the number of modulations changes systematically between notes, with an increase or a decrease of one modulation only. Adherence to these rules is exceedingly robust. Of the 2536 calls analyzed, 100% began with a note containing no modulations, 91.5% followed the “modulation begins with one” rule, and 97.5% followed the “plus or minus one modulation” rule. This pattern is consistent between individual males, as well as across choruses and breeding sites. The regularity of the pattern suggests that it serves some biological significance in the communication behavior of the bullfrog.

A. Markov chains, properties, and processes

Markov chain analyses were developed as a way of determining if events that occur in sequence are independent of one another. A crucial feature of these analyses is the concept of uncertainty (or entropy), which is calculated using the number of events that occur in a particular sequence. When such events are completely independent of one another (i.e., the previous event has absolutely no influence on the immediately following event), the data are described as exhibiting the Markov property. At the other end of the spectrum are previously occurring events that absolutely determine the appearance of the following event. These types of data are termed Markov processes. In a Markov process, knowing which event occurs first will completely predict which event will follow. Animal behavior often falls between these two

extremes and is therefore labeled semi-Markovian (Fagen and Young, 1978). This simply means that the last event the animal engages in may be a good predictor, albeit not a perfect one, of the next event to follow. Because different sets of data vary in their degree of approach to true Markov processes, information theory (Shannon and Weaver, 1949) can be used to estimate how much uncertainty occurs in a sequence of events, and thus provide an estimate of the amount of redundancy or diversity in a particular data set. This kind of analysis has become a powerful quantitative tool for examining and analyzing animal communication systems in diverse species and across a range of modalities. Information theory has been used to categorize the whistle signals of dolphins (McGowan *et al.*, 1999, 2002) and the call repertoires of Mexican chickadees (Ficken *et al.*, 1993), and to analyze aggressive communications in mantis shrimp (Dingle, 1969) and hermit crabs (Hazlett and Bossert, 1965). It has also been applied to analysis of intermale behavior of grasshoppers (Steinberg and Conant, 1974), the study of brood recognition in Mexican free-tailed bats (Beecher, 1989), and the mutually symbiotic relationship between goby fish and snapping shrimp (Preston, 1978). It is this generality of application that makes information theory so useful in examining the capacity, structure, and organization of animal repertoires.

Uncertainty analyses reveal that the appearance of modulations in bullfrog advertisement calls follows a semi-Markovian, deterministic pattern. This is shown in the sizable drop in uncertainty that occurs between U_0 and U_1 , with less of a drop between U_1 and U_2 . The drop between U_m and U_0 shows that knowing the relative frequencies of the seven events in the data set has decreased the amount of uncertainty, and the drop from U_0 to U_1 has established the existence of a pattern. The low value of U_1 is consistent with the “plus or minus 1” rule, since knowing how many modulations (x) the first note has strongly implies the next note will contain x , $x + 1$, or $x - 1$ modulations. The analyses suggest that information transfer is occurring based upon the ordered frequencies between successive events.

B. Mechanisms of production

Calling is an energetically costly behavior for frogs, with an oxygen demand that exceeds that occurring during forced exercise (Taigen and Wells, 1985). In *Hyla*, the trunk muscles are specialized for high-power output, but are highly susceptible to fatigue (Girgenrath and Marsh, 1999). It is possible that the pattern of insertion of modulations reported here reflects fatigue of the vocal musculature. For example, the production of modulations in a series of successive notes might indicate that the male was simply running out of breath. Such a fatigue hypothesis implies that bullfrogs would incorporate more modulations in their calls at the end than at the beginning of a nightly chorus, and that the aggregate chorus would show more modulations at the end than near the beginning of the calling season. Our behavioral data do not show such a simple trend.

In *Bufo*, amplitude modulation of vocalizations is produced by both active (pattern of contraction of the thoracic musculature) and passive (vibration of the arytenoid carti-

lages) mechanisms (Martin, 1971). Which mechanism best accounts for the pattern of modulation seen here cannot be determined from the present data. One possibility is that the production of longer call notes might passively engage the arytenoid cartilages, thus producing modulations. The positive relationship between note duration and number of modulations supports this hypothesis. Inserting modulations might also allow the male to extend the length of the individual notes in his calls while reducing the amount of air exchange needed. This in turn implies that longer duration notes have a different perceptual significance than shorter duration ones. Male bullfrogs do not respond readily to playbacks of short-duration notes outside the normal range of variability in note duration, but they respond equally vigorously to playbacks of modal duration and long-duration notes, showing patterns of generalization rather than discrimination for this acoustic feature (Simmons, 2004). These data suggest that the insertion of modulations *per se* might not have a perceptual significance to other males in the chorus. On the other hand, female bullfrogs, like female frogs of other species (Gerhardt and Huber, 2002), may prefer notes with longer rather than shorter durations, and use note duration as an indicator of male quality. This hypothesis has not been tested directly.

C. Physiological relevance of envelope modulation

Envelope periodicity is an important feature mediating vocal interactions in many anuran species (Gerhardt and Huber, 2002). In playback experiments, male bullfrogs respond differentially to synthetic advertisement calls with unmodulated and modulated notes, when these envelope differences are produced by changes in either the phase structure or the harmonic structure of the signal (Hainfeld *et al.*, 1996; Simmons and Bean, 2000). Although these experiments did not explicitly test responses to signals that resemble the modulation pattern described here, they predict that modulation is a salient perceptual cue for these animals. Physiological experiments show that the 100-Hz periodicity of unmodulated synthetic notes and the envelope fluctuations produced by changes in phase structure are encoded in the bullfrog's eighth nerve by a neural code based on phase-locked discharges (Schwartz and Simmons, 1990; Simmons *et al.*, 2001). This temporal code seems to weaken in the central auditory system, such that, in the auditory midbrain (torus semicircularis, TS), only a relatively small number of neurons respond with phase-locked discharges to modulations in the 90–100-Hz range (Simmons *et al.*, 2000). Yet, strong phase-locked activity continues to occur even in the TS in response to signals with amplitude modulations of 20–50 Hz (Bibikov and Nizamov, 1996; Simmons *et al.*, 2000). The modulated notes described in this study have low-frequency modulations of 12 Hz or less. Although such low frequencies have not been tested in studies of modulation coding in the frog's auditory midbrain, we predict that these stimuli would produce strong phase locking, thus retaining a temporal code for processing of the advertisement call in the central auditory system.

ACKNOWLEDGMENTS

This research was supported by NIH Grant DC05257 to A.M.S. and a Brown University Undergraduate Teaching and Research Award to D.N.S. We thank Seth Horowitz and two anonymous reviewers for their helpful comments on the manuscript.

- Bee, M. A. (2004). "Within-individual variation in bullfrog vocalizations: Implications for a vocally mediated social recognition system," *J. Acoust. Soc. Am.* **116**, 3770–3781.
- Bee, M. A., and Gerhardt, H. C. (2001). "Neighbor–stranger discrimination by territorial male bullfrogs (*Rana catesbeiana*). I. Acoustic basis," *Anim. Behav.* **62**, 1129–1140.
- Beecher, M. D. (1980). "Signaling systems for individual recognition: An information theory approach," *Anim. Behav.* **38**, 248–261.
- Bibikov, N. G., and Nizamov, S. V. (1996). "Temporal coding of low-frequency amplitude modulation in the torus semicircularis of the grass frog," *Hear. Res.* **101**, 23–44.
- Boatright-Horowitz, S. L., Horowitz, S. S., and Simmons, A. M. (2000). "Patterns of vocal response in a bullfrog (*Rana catesbeiana*) chorus: Preferential responding to far neighbors," *Ethology* **106**, 701–712.
- Capranica, R. R. (1965). *The Evoked Vocal Response of the Bullfrog: A Study of Communication by Sound* (MIT Press: Cambridge, MA).
- Davis, M. S. (1987). "Acoustically mediated neighbor recognition in the North American bullfrog, *Rana catesbeiana*," *Behav. Ecol. Sociobiol.* **21**, 185–190.
- Dingle, H. (1969). "A statistical and information analysis of aggressive communication in the mantis shrimp, *Gonodactylus bredini* Manning," *Anim. Behav.* **17**, 561–575.
- Emlen, S. T. (1968). "Territoriality in the bullfrog, *Rana catesbeiana*," *Copeia* **2**, 240–243.
- Fagen, R. M., and Young, D. Y. (1978). "Temporal patterns of behavior: Durations, intervals, latencies, and sequences," in *Quantitative Ethology*, edited by P. W. Colgan (Wiley, New York).
- Ficken, M. S., Hailman, E. D., and Hailman, J. P. (1994). "The chick-a-dee call system of the Mexican chickadee," *Condor* **96**, 70–82.
- Gerhardt, H. C., and Huber, F. (2002). *Acoustic Communication in Insects and Anurans: Common Problems and Diverse Solutions* (University Chicago Press, Chicago, IL).
- Girgenrath, M., and Marsh, R. L. (1999). "Power output of sound-producing muscles in the tree frogs *Hyla versicolor* and *Hyla chrysoscelis*," *J. Exp. Biol.* **202**, 3225–3237.
- Hailman, E. D., and Hailman, J. P. (1993). *UNCERT User's Guide* (University of Wisconsin Zoology Department, Madison, WI). Available at (<http://www.animalbehavior.org/Resources/CSASAB/#Uncert>)
- Hainfeld, C. A., Boatright-Horowitz, S. L., Boatright-Horowitz, S. S., and Simmons, A. M. (1996). "Discrimination of phase spectra in complex sounds by the bullfrog (*Rana catesbeiana*)," *J. Comp. Physiol. A* **179**, 75–87.
- Hazlett, B., and Bossert, W. (1965). "A statistical analysis of the aggressive communication systems of some hermit crabs," *Anim. Behav.* **13**, 357–373.
- Lewis, D. B., and Gower, D. M. (1980). *Biology of Communication* (Wiley, New York).
- Martin, W. F. (1971). "Mechanics of sound production in toads of the genus *Bufo*: Passive elements," *J. Exp. Zool.* **176**, 273–293.
- McGowan, B., Hanser, S. F., and Doyle, L. R. (1999). "Quantitative tools for comparing animal communication systems: Information theory applied to bottlenose dolphin whistle repertoires," *Anim. Behav.* **57**, 409–419.
- McGowan, B., Hanser, S. F., and Doyle, L. R. (2002). "Using information theory to assess the diversity, complexity, and development of communicative repertoires," *J. Comp. Psychol.* **116**, 166–172.
- Narins, P. M. and Capranica, R. R. (1978). "Communicative significance of the two note call of the treefrog, *Eleutherodactylus coqui*," *J. Comp. Physiol.* **127**, 1–9.
- Preston, J. L. (1978). "Communication systems and social interactions in a goby-shrimp symbiosis," *Anim. Behav.* **26**, 791–802.
- Schwartz, J. J., and Simmons, A. M. (1990). "Encoding of a spectrally complex communication sound in the bullfrog's auditory nerve," *J. Comp. Physiol., A* **166**, 489–500.
- Shannon, C. E., and Weaver, W. (1949). *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, IL).

- Simmons, A. M. (2004). "Call recognition in the bullfrog, *Rana catesbeiana*: Generalization along the duration continuum," J. Acoust. Soc. Am. **115**, 1345–1355.
- Simmons, A. M., and Bean, M. E. (2000). "Perception of mistuned harmonics in complex sounds by the bullfrog (*Rana catesbeiana*)," J. Comp. Psychol. **114**, 167–173.
- Simmons, A. M., Sanderson, M. I., and Garabedian, C. E. (2000). "Representation of waveform periodicity in the auditory midbrain of the bullfrog, *Rana catesbeiana*," J. Assoc. Res. Otolaryn. **1**, 2–24.
- Simmons, A. M., Eastman, K., and Simmons, J. A. (2001). "Autocorrelation model of periodicity coding in bullfrog auditory nerve fibers," ARLO **2**, 1–6.
- Steinberg, J. B., and Conant, R. D. (1974). "An informational analysis of the inter-male behavior of the grasshopper *Chortophaga viridifasciata*," Anim. Behav. **22**, 617–627.
- Sustare, D. (1978). "Systems diagrams," in *Quantitative Ethology*, edited by P. W. Colgan (Wiley, New York).
- Taigen, T. L., and Wells, K. D. (1985). "Energetics of vocalization by an anuran amphibian (*Hyla versicolor*)," J. Comp. Physiol. B **155**, 163–170.
- Wells, K. D. (1977). "The social behavior of anuran amphibians," Anim. Behav. **2**, 666–693.
- Wiewandt, T. A. (1969). "Vocalization, aggressive behavior, and territoriality in the bullfrog, *Rana catesbeiana*," Copeia **2**, 276–284.

Gas bubble and solid sphere motion in elastic media in response to acoustic radiation force

Yurii A. Ilinskii, G. Douglas Meegan, and Evgenia A. Zabolotskaya

Applied Research Laboratories, The University of Texas at Austin, Austin, Texas 78713-8029

Stanislav Y. Emelianov

Biomedical Engineering Department, The University of Texas at Austin, Austin, Texas 78712-1084

(Received 6 August 2004; revised 6 January 2005; accepted 11 January 2005)

The general approach to estimate the displacement of rounded objects (specifically, gas bubbles and solid spheres) in elastic incompressible media in response to applied acoustic radiation force is presented. In this study, both static displacement and transient motion are analyzed using the linear approximation. To evaluate the static displacement of the spherical inclusion, equations coupling the applied force, displacement, and shear modulus of the elastic medium are derived. Analytical expressions to estimate the static displacement of solid spheres and gas bubbles are presented. Under a continuously applied static force, both the solid sphere and the initially spherical gas bubble are displaced, and the bubble is deformed. The transient responses of the inclusions are described using motion equations. The displacements of the inclusion in elastic incompressible lossless media are analyzed using both frequency-domain and time-domain formalism, and the equations of motion are derived for both a solid sphere and a gas bubble. For a short pulsed force, an analytical solution for the equations of motion is presented. Finally, transient displacement of the gas bubble in viscoelastic media is considered. © 2005 Acoustical Society of America.

[DOI: 10.1121/1.1863672]

PACS numbers: 43.80.Qf [FD]

Pages: 2338–2346

I. INTRODUCTION

Recent research on acoustic radiation force has shown great promise for several applications including biomedical ultrasound. For example, shear waves may be induced in tissue by the radiation force of a focused ultrasonic beam and provide shear wave imaging.¹ Another application offers remote palpation using a radiation force method to get information about the local mechanical properties of tissue.^{2–4} In these applications the acoustic radiation force generates localized tissue displacement that depends on its mechanical properties. The radiation force in tissue-like media is determined by the absorption and reflection of ultrasound from inhomogeneities. Lesions or other targets (like gas bubbles or solid spheres) can be considered as inhomogeneities.

A few studies utilizing the radiation force acting on a gas bubble⁵ and a solid sphere⁶ have been reported recently. Erpelding *et al.*⁵ described how the acoustic radiation force was applied to a gas bubble (generated by femtosecond laser pulses focused in the volume of tissue-mimicking gels). Bubble displacements in gelatin phantoms of different concentrations were measured and the maximum displacement reached was several hundreds of micrometers.⁵ Furthermore, it was shown that the bubble displacement depended on the elastic (Young's or shear) modulus of the gel.

In a study by Chen *et al.*, the response of a solid sphere, embedded into an otherwise homogeneous gelatin phantom, was measured.⁶ The mechanical excitation of the sphere was produced using a low-frequency radiation force. The velocities of vibrating spheres of different radii were measured in both liquid (water) and elastic (gelatin) medium. Calculations were performed using the analytical model based on

Oestreicher's theory as developed for the oscillations of a rigid sphere in a viscoelastic medium.⁷ The results of the study demonstrate that mechanical response of the sphere under radiation force is related to viscoelastic properties of the surrounding medium and can be used to determine the local material properties.⁶

Three methods to estimate the radiation force \mathbf{F}_r in acoustic field are often used. The first method is based on the equation derived for a bubble pulsating in liquid^{8,9}

$$\mathbf{F}_r = -\langle V \nabla p_a \rangle, \quad (1)$$

where V is the bubble volume, ∇p_a is a gradient of the acoustic pressure in a medium, and $\langle \rangle$ denotes an average with respect to time.

The second method commonly used is based on conservation of momentum.¹⁰ Assuming that the bubble radius is much smaller than an acoustic wavelength, the scattered wave is nearly spherical and does not carry away a momentum, so the momentum of the incident wave is changed due to absorption and bubble radiation by the amount W_r/c_0 per unit time (W_r is the total radiated and absorbed acoustic power, c_0 is the speed of sound.) Neglecting the absorption and applying the equation for the pressure in far field $p_r = \rho \ddot{V}/4\pi r$ presented in Ref. 11, the radiated power yields

$$W_r = \frac{4\pi r^2}{\rho c_0} p_r^2 = \frac{\rho}{4\pi c_0} \dot{V}^2, \quad (2)$$

where \ddot{V} is the second derivative of the bubble volume with respect to time. It is worth noting that Eq. (2) can be applied to bubble oscillations of both small and finite amplitude.¹¹ The acoustic radiation force is

$$F_r = \frac{W_r}{c_0} = \frac{\rho}{4\pi c_0^2} \langle \ddot{V}^2 \rangle. \quad (3)$$

The third method reviewed here is as follows. The radiation pressure on the target can be estimated by integrating the acoustic field in the surrounding medium with respect to the target surface. For example, Hasegawa and Yosioka used this method to calculate the radiation force on a solid elastic sphere in a plane progressive sound field in liquid.¹²

These approaches have been developed to evaluate the radiation force in liquids, but they can also be applied to estimate the radiation force in tissue-like media. Rigorous verification of the last statement will be presented in our next publication. In the analysis presented here, the radiation force is assumed to be given.

In this paper, the static displacement and transient motion of a gas bubble and a rigid sphere in a viscoelastic medium is estimated and discussed. The current study is focused on developing a theoretical model to describe both static and transient displacements of a target in soft elastic media in response to the externally applied radiation force. The analysis is performed in the linear approximation.

The static displacements of a sphere and a gas bubble in elastic media are considered first. In this case, the pushing force is considered to be given and constant, and the elastic medium is assumed to be isotropic, homogeneous, incompressible, and inviscous. A technique similar to the one applied to estimate Stokes' drag of a solid sphere in viscous incompressible liquid is used here. The transient motions of the sphere and the gas bubble in the elastic, incompressible, inviscous medium are investigated in response to a time-varying force. Both monochromatic and impulse forces are considered. Finally, the transient motion of the bubble in the viscoelastic incompressible media is analyzed. The paper concludes with discussion of the results. Results of this study were partly presented at the 147th Meeting of the Acoustical Society of America.¹³

II. STATIC DISPLACEMENT

Static displacement of an object that can be a rigid sphere or a gas bubble in an elastic medium is estimated in the linear approximation. The applied force is assumed to be constant, i.e., not varying with time. While in a liquid the object is displaced as long as the force is applied, in an elastic medium the displacement is constrained by the elastic properties of the medium. In our analysis the elastic medium is assumed to be isotropic, homogeneous, incompressible, and inviscous. The technique used here is similar to Stokes' formula derivation for the drag of a solid sphere in viscous, incompressible liquid.

For small Reynolds numbers, the Stokes' force, F_s , acting on a solid sphere moving with velocity v in viscous incompressible liquid is (see, for example, Ref. 14)

$$F_s = 6\pi\eta Rv, \quad (4)$$

where R is a solid sphere radius, and η is the viscosity coefficient of liquid. The starting point of Stokes' formula deri-

vation is the equation for the velocity of incompressible viscous liquid¹⁴

$$\eta\nabla^2\mathbf{v} = \nabla p, \quad (5)$$

and the continuity equation

$$\nabla \cdot \mathbf{v} = 0, \quad (6)$$

where \mathbf{v} is a relative velocity between the sphere and surrounding liquid, and p is the pressure in the liquid.

In the elastic medium only finite displacement of the bubble or the sphere from an initial position takes place. The balance equation for stress tensor σ_{ik} in an incompressible isotropic elastic medium is^{15,16}

$$\frac{\partial \sigma_{ik}}{\partial x_k} = 0. \quad (7)$$

Taking into account that the stress tensor in incompressible elastic medium is

$$\sigma_{ik} = -p\delta_{ik} + 2\mu u_{ik}, \quad (8)$$

and the strain tensor u_{ik} in linear approximation is

$$u_{ik} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} \right), \quad (9)$$

Eq. (7) can be reduced to the equation

$$\mu\nabla^2\mathbf{u} = \nabla p. \quad (10)$$

Here, \mathbf{u} is a displacement vector, μ is the shear modulus of the elastic incompressible medium, and p is the isotropic part of a stress tensor. Incompressibility of the elastic medium in the linear approximation requires

$$\nabla \cdot \mathbf{u} = 0. \quad (11)$$

Equations (10) and (11) for displacement in an elastic incompressible medium are identical to Eqs. (5) and (6) for velocity in viscous incompressible liquid. Therefore, the technique developed to solve Eqs. (5) and (6), presented in detail in Ref. 14, can be used to solve Eqs. (10) and (11).

The solution of Eqs. (10) and (11) can be presented in the form

$$\mathbf{u} = \nabla \times (\nabla f \times \mathbf{e}) = \nabla \times \nabla \times (f\mathbf{e}). \quad (12)$$

Here, $f = f(r)$ is a scalar function

$$f(r) = ar + \frac{b}{r}, \quad (13)$$

and \mathbf{e} is a unit vector along the displacement, and constants a and b are determined by boundary conditions on the object surface.

If the polar axis of the spherical system of coordinates is along the displacement vector (i.e., an angle θ is between a radius vector and displacement), the equations for the components of the displacement vector \mathbf{u} are

$$u_r = \frac{2}{r^2} \left(-ar + \frac{b}{r} \right) \cos \theta, \quad u_\theta = \frac{1}{r} \left(a + \frac{b}{r^2} \right) \sin \theta. \quad (14)$$

From Eq. (10) the pressure gradient created by the displaced object is

$$\nabla p = \mu \nabla^2 \mathbf{u} = \mu \nabla^2 [\nabla \times \nabla \times (f\mathbf{e})]. \quad (15)$$

Following the derivations presented in Ref. 14, the pressure is

$$p = p_0 + \mu \nabla^2 [\nabla \cdot (f\mathbf{e})] = p_0 + \mu (\mathbf{e} \cdot \nabla) \nabla^2 f, \quad (16)$$

where p_0 is an integration constant that describes the pressure far away from the object. The force on the displaced solid sphere or bubble in the elastic medium is¹⁴

$$F = \int (-p \cos \theta + \sigma_{rr} \cos \theta - \sigma_{r\theta} \sin \theta) dS. \quad (17)$$

Here, dS is an element of the sphere surface, and σ_{rr} and $\sigma_{r\theta}$ are components of a stress tensor given by

$$\sigma_{ik} = \mu \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} \right). \quad (18)$$

Due to the symmetry of the problem, Eq. (17) should be integrated with respect to θ only

$$F = 2\pi R^2 \int_0^\pi (-p \cos \theta + \sigma_{rr} \cos \theta - \sigma_{r\theta} \sin \theta) \sin \theta d\theta. \quad (19)$$

Consequently, the stress tensor components written in the spherical system of coordinates are

$$\sigma_{r\theta} = 2\mu u_{r\theta} = \mu \left(\frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} + \frac{1}{r} \frac{\partial u_r}{\partial \theta} \right), \quad (20)$$

$$\sigma_{rr} = 2\mu u_{rr} = 2\mu \frac{\partial u_r}{\partial r}. \quad (21)$$

In the following sections, the displacements of the solid sphere and the gas bubble in the elastic medium are considered.

A. Static displacement of the solid sphere

The boundary conditions at the surface of the rigid solid sphere of radius R displaced by a vector \mathbf{u}_0

$$u_r(R, \theta) = u_0 \cos \theta, \quad u_\theta(R, \theta) = -u_0 \sin \theta, \quad (22)$$

allow us to find the constants a and b

$$a = -\frac{3Ru_0}{4}, \quad b = -\frac{R^3u_0}{4}. \quad (23)$$

Therefore, Eqs. (14) can be rewritten

$$u_r(r, \theta) = u_0 \left(\frac{3R}{2r} - \frac{R^3}{2r^3} \right) \cos \theta, \quad (24)$$

$$u_\theta(r, \theta) = -u_0 \left(\frac{3R}{4r} + \frac{R^3}{4r^3} \right) \sin \theta.$$

The function f in this case is

$$f = ar + \frac{b}{r} = -\frac{Ru_0}{4} \left(3r + \frac{R^2}{r} \right). \quad (25)$$

Calculating $\nabla^2 f$ and substituting into Eq. (16) yields

$$p = p_0 - \frac{3\mu Ru_0}{2} (\mathbf{e} \cdot \nabla) \frac{1}{r} = p_0 + \frac{3\mu Ru_0}{2r^2} \cos \theta. \quad (26)$$

The necessary stress tensor components are then

$$\sigma_{r\theta} = \frac{3\mu u_0}{2r^4} R^3 \sin \theta, \quad (27)$$

$$\sigma_{rr} = -\frac{3\mu u_0}{r} \left(\frac{R}{r} - \frac{R^3}{r^3} \right) \cos \theta. \quad (28)$$

On the surface of the sphere ($r=R$), the radial component of the stress vanishes ($\sigma_{rr}=0$). Substituting Eqs. (26) and (27) into Eq. (19) and integrating yields

$$F = -6\pi\mu Ru_0. \quad (29)$$

Equation (29) is similar to Stokes' formula presented here by Eq. (4), where the viscous coefficient η is replaced by the shear modulus μ , and velocity v is replaced by the displacement u_0 . Equation (29) accounts for the force acting on the spherical particle displaced by the vector \mathbf{u}_0 in the elastic medium. If an external force (i.e., the radiation force) is applied to the sphere, the sign should be changed

$$F_r = 6\pi\mu Ru_0, \quad (30)$$

and the displacement of the sphere is

$$u_0 = \frac{F_r}{6\pi\mu R}. \quad (31)$$

Therefore, the displacement of the solid sphere in the elastic medium is proportional to the applied force and inversely proportional to the shear modulus of the medium and the radius of the sphere.

Equation (31) can be applied to estimate displacements of nonspherical targets.

B. Static displacement of a gas bubble

Static displacement of a gas bubble in the elastic medium is estimated with the same equations but the boundary condition is different. The boundary condition at the surface of the gas bubble ($r=R$) is

$$\sigma_{r\theta} = 0. \quad (32)$$

The stress tensor component $\sigma_{r\theta}$ can be evaluated using Eq. (20), where u_r and u_θ are given by Eqs. (14)

$$\sigma_{r\theta} = \mu \left(\frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} + \frac{1}{r} \frac{\partial u_r}{\partial \theta} \right) = -\frac{6\mu}{r^4} b \sin \theta. \quad (33)$$

The boundary condition requires that $b=0$. Therefore, the components of the displacement vector are

$$u_r = -\frac{2a}{r} \cos \theta, \quad u_\theta = \frac{a}{r} \sin \theta, \quad (34)$$

and $\nabla^2 f$ is

$$\nabla^2 f = \frac{2a}{r}. \quad (35)$$

The pressure is calculated using Eq. (16)

$$p = p_0 + 2\mu a(\mathbf{e} \cdot \nabla) \frac{1}{r} = p_0 - \frac{2\mu a}{r^2} \cos \theta. \quad (36)$$

The component of the stress tensor σ_{rr} is

$$\sigma_{rr} = 2\mu \frac{\partial u_r}{\partial r} = \frac{4\mu a}{r^2} \cos \theta. \quad (37)$$

The value a in Eq. (37) is estimated using the equilibrium equation

$$p - \sigma_{rr} + p_r = p_g, \quad (38)$$

when pressure outside the bubble (including the pressure in the medium p , the normal stress σ_{rr} acting on the bubble, and the acoustic radiation pressure p_r) is in equilibrium with the internal gas pressure in the bubble p_g . The radiation pressure p_r here is the external force per unit area acting on the bubble surface. Since the radiation force changes with θ , the pressure p_r is also a function of θ

$$p_r = p_r^{(0)} \cos \theta, \quad (39)$$

where $p_r^{(0)}$ is the pressure for $\theta=0$, and its value is defined by the equation for the radiation force

$$F_r = 2\pi R^2 \int_0^\pi (-p_r \cos \theta) \sin \theta d\theta. \quad (40)$$

Integration of Eq. (40) yields

$$p_r^{(0)} = -\frac{3F_r}{4\pi R^2}. \quad (41)$$

Substituting Eqs. (36), (37), (39), and (41) into Eq. (38) and setting $r=R$ as well as $p_0=p_g$ gives

$$a = -\frac{F_r}{8\pi\mu}. \quad (42)$$

The displacement components of the bubble surface are then given by

$$\begin{aligned} u_r &= \frac{F_r}{4\pi\mu R} \cos \theta = \alpha R \cos \theta, \\ u_\theta &= -\frac{F_r}{8\pi\mu R} \sin \theta = -\frac{\alpha R}{2} \sin \theta, \end{aligned} \quad (43)$$

where new notation α is introduced

$$\alpha = \frac{F_r}{4\pi\mu R^2}. \quad (44)$$

The dimensionless parameter α is the ratio between the externally applied force and the value associated with elastic property of the medium.

It is easy to show that the bubble changes its shape during displacement. Indeed, the initial shape of the bubble is assumed to be spherical and, in Cartesian coordinates, can be expressed as

$$x = R \cos \theta, \quad y = R \sin \theta. \quad (45)$$

Here, x coincides with a polar axis of a spherical coordinate system and, therefore, the angle θ is an angle from the x axis.

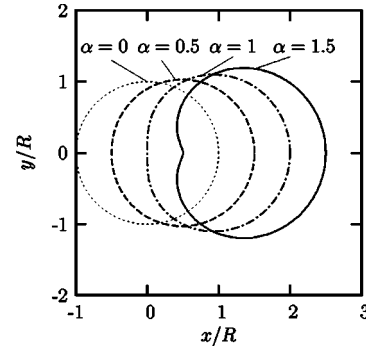


FIG. 1. Displacements and shapes of the bubble for different ratio of the applied force and medium stiffness: the parameter $\alpha = F_r/4\pi\mu R^2$ takes values 0; 0.5; 1; 1.5. The curve for $\alpha=0$ is initial bubble shape and position: the bubble center is located at $x=0$, $y=0$. The curve for $\alpha=0.5$ corresponds to the relatively weak force, the bubble is displaced and remains almost spherical. The curve for $\alpha=1$ is $F_r=4\pi\mu R^2$; the bubble is displaced and loses the spherical form. The curve for $\alpha=1.5$ corresponds to the relatively strong force; the bubble displacement is greater and its deformation is much stronger.

The relationship between Cartesian and spherical coordinates is

$$x = r \cos \theta, \quad y = r \sin \theta. \quad (46)$$

During the bubble displacement, the coordinates x and y changes incrementally by δx and δy

$$\delta x = \delta r \cos \theta - r \sin \theta \delta \theta, \quad \delta y = \delta r \sin \theta + r \cos \theta \delta \theta. \quad (47)$$

Taking into account that

$$\delta r = u_r, \quad r \delta \theta = u_\theta, \quad (48)$$

the shape of the displaced bubble can be presented in the form

$$\frac{x}{R} = \cos \theta + \frac{\alpha}{4} (3 + \cos 2\theta), \quad \frac{y}{R} = \sin \theta + \frac{\alpha}{4} \sin 2\theta. \quad (49)$$

For $\theta=0$, Eqs. (49) yield

$$\frac{x}{R} = 1 + \alpha. \quad (50)$$

It is clear from Eq. (50) that parameter α has a simple physical interpretation—it characterizes the ratio between the displacement of the bubble surface at $\theta=0$ and the bubble radius.

Displacements and shapes of the bubble for the different forces (different parameters α) are plotted in Fig. 1. As expected, the bubble displacement becomes greater with increasing the parameter α , and for $\alpha \geq 1$ bubble deformation is getting noticeable. Although the results presented in Fig. 1 are reasonable, they should be used with caution. The problem has been formulated in linear approximation, assuming that the displacements are small compared to the bubble radius.

Since the influence of surface tension is pretty clear and the procedure to take it into account is known, we have not considered the surface tension here and note only that the surface tension reduces the effect described above, especially

for small bubbles. Surface tension makes the small bubble behave like a solid particle. The influence of surface tension on the bubble displacement and its shape will be discussed in detail in our future publications.

III. TRANSIENT DISPLACEMENT

We will now consider the displacements of solid sphere and gas bubble in the elastic incompressible medium when the external force depends on time. Since the problem is analyzed in linear approximation, force dependence on time is taken as monochromatic, i.e., proportional to $\mathbf{e}^{-i\omega t}$.

Again, the technique used here is very close to the method presented in Ref. 14 where the oscillatory motion of a sphere in viscous incompressible liquid is considered, and the drag force is estimated. If the periodic force is applied to the object, then its movement has an oscillatory character. The equation for sphere motion in the elastic medium is

$$-\rho\omega^2\mathbf{u} = \mu\nabla^2\mathbf{u} - \nabla p, \quad (51)$$

where ρ is elastic medium density and ω is a frequency of sphere oscillations induced by monochromatic force.

The solution of Eq. (51) is formally given by Eq. (12), where the function $f(r)$ is defined as

$$f(r) = \frac{\tilde{a}\mathbf{e}^{ikr}}{ikr} - \frac{\tilde{b}}{r}. \quad (52)$$

Here, \tilde{a} , \tilde{b} are complex constants that are determined by the boundary conditions and k is a wave number for a shear wave of frequency ω

$$k^2 = \frac{\omega^2}{c_t^2} = \frac{\rho\omega^2}{\mu}. \quad (53)$$

The components of the displacement vector \mathbf{u} in this case are

$$u_r = -\frac{2}{r^2} \left[\tilde{a}\mathbf{e}^{ikr} \left(1 - \frac{1}{ikr} \right) + \frac{\tilde{b}}{r} \right] \cos \theta, \quad (54)$$

$$u_\theta = -\frac{1}{r^2} \left[\tilde{a}\mathbf{e}^{ikr} \left(1 - ikr - \frac{1}{ikr} \right) + \frac{\tilde{b}}{r} \right] \sin \theta. \quad (55)$$

The pressure gradient is determined by

$$\nabla p = \mu\nabla[\nabla^2 + k^2](\nabla \cdot \mathbf{f}\mathbf{e}), \quad (56)$$

and the equation for pressure is

$$p = p_0 + \mu(\mathbf{e} \cdot \nabla)(\nabla^2 f + k^2 f), \quad (57)$$

where p_0 is an integration constant corresponding to pressure far away from the object.

A. Transient displacement of a solid sphere

The boundary conditions for the solid sphere are the same as in the static case and are given by Eq. (22). The constants \tilde{a} and \tilde{b} calculated using Eq. (22) are

$$\tilde{a} = -\frac{3Ru_\omega}{2ik}\mathbf{e}^{-ikR}, \quad \tilde{b} = -\frac{R^3u_\omega}{2} \left(1 - \frac{3}{ikR} - \frac{3}{k^2R^2} \right), \quad (58)$$

where u_ω are spectral components of a displacement.

Equation (52) for the function f allows one to calculate $\nabla^2 f + k^2 f$, that is

$$\nabla^2 f + k^2 f = -\frac{k^2\tilde{b}}{r}. \quad (59)$$

Substitution of Eq. (59) into Eq. (57) and elimination of \tilde{b} using Eqs. (58) yields

$$p = p_0 - \frac{\mu k^2 R^3 u_\omega}{2r^2} \left(1 - \frac{3}{ikR} - \frac{3}{k^2 R^2} \right) \cos \theta. \quad (60)$$

The components of the stress tensor on the surface of the sphere are

$$\sigma_{r\theta} = \frac{3\mu u_\omega}{2R} (1 - ikR) \sin \theta, \quad \sigma_{rr} = 0. \quad (61)$$

Ultimately, spectral components of the force calculated using Eq. (19) are

$$F_\omega = -6\pi\mu R u_\omega (1 - ikR - \frac{1}{9}k^2 R^2). \quad (62)$$

The equation that couples the solid sphere displacement and the external periodic force $F_\omega^{(\text{ext})}$, for example, the radiation force, is

$$\begin{aligned} F_\omega^{(\text{ext})} &= -M\omega^2 u_\omega - F_\omega \\ &= -M\omega^2 u_\omega + 6\pi\mu R u_\omega (1 - ikR - \frac{1}{9}k^2 R^2), \end{aligned} \quad (63)$$

where M is a mass of the solid sphere.

Equation (63) has simple interpretation: it is an equation for solid sphere motion in the elastic incompressible medium written in frequency domain. Indeed, the term $6\pi R^2 \mu u_\omega ik$ takes into account damping of sphere's oscillations due to radiation of shear waves, and the term $2\pi R^3 \mu u_\omega k^2/3$ accounts for motion of the medium around the sphere. As $k = \omega/c_t$ and $c_t = (\mu/\rho)^{1/2}$, the last term in Eq. (63) can be written in the form

$$-\frac{2}{3}\pi R^3 \mu u_\omega k^2 = -\frac{2}{3}\pi R^3 \rho \omega^2 u_\omega = -m\omega^2 u_\omega. \quad (64)$$

Here, $-\omega^2 u_\omega$ is the spectral presentation of the second derivative of the displacement $u(t)$, and m , the so-called induced mass

$$m = \frac{2}{3}\pi R^3 \rho, \quad (65)$$

determines the mass of the elastic medium involved in motion. It is worth noting that the induced mass is equal to half of the sphere mass with radius R and density ρ . The term $-6\pi R^2 \mu u_\omega ik$ in Eq. (63) also can be presented in the form

$$-6\pi R^2 \mu u_\omega ik = -6\pi R^2 \mu c_t^{-1} i\omega u_\omega, \quad (66)$$

where $-i\omega u_\omega$ is the spectral presentation of the first derivative of the displacement. Equation (66) describes the dampening force that is proportional to the velocity of the solid sphere.

Equation (63) for solid sphere motion can be written in the time domain. In dimensionless variables

$$U = \frac{u(t)}{R}, \quad \tilde{F}_{\text{ext}} = \frac{F_{\text{ext}}(t)}{6\pi\mu R^2}, \quad (67)$$

the time-domain equation is

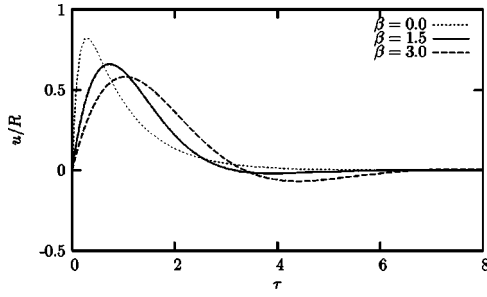


FIG. 2. Responses of solid spheres with different densities (the parameter β takes values 0; 1.5; 3) to the short pulsed force calculated with Eq. (68): the external force is taken as a δ function, $\tau = tc_t/R$ is dimensionless time.

$$U + \dot{U} + \frac{1}{9}(1 + 2\beta)\ddot{U} = \tilde{F}_{\text{ext}}, \quad (68)$$

where dots above the function U denote the derivatives with respect to dimensionless time τ , and parameter β is solid sphere density ρ_s normalized by medium density

$$\tau = \frac{tc_t}{R}, \quad \beta = \frac{\rho_s}{\rho}. \quad (69)$$

The equations for solid sphere displacement in the elastic medium—Eq. (63) in the frequency domain and Eq. (68) in the time domain—can be solved analytically or numerically for any external force applied to the sphere.

The responses of the solid sphere with different densities to the short pulsed force taken as a δ function have been calculated with Eq. (68). The results are presented in Fig. 2. There are no oscillatory displacements for the empty sphere $M=0$ and for the relatively light sphere, while the heavier spheres may exhibit a few oscillations. The empty sphere is included here for late comparison with gas bubble.

Results presented above have been obtained for incompressible elastic media. Compressibility of the elastic medium can be also included. In this case, the equation for sphere motion is

$$-\rho\omega^2\mathbf{u} = \mu\nabla^2\mathbf{u} + (\lambda + \mu)\nabla(\nabla \cdot \mathbf{u}), \quad (70)$$

where λ is a Lamé parameter. Equation (70) replaces Eq. (51), which has been used for incompressible media. A solution of Eq. (70) consists of two parts: an incompressible one

$$-\rho\omega^2\mathbf{u}_s = \mu\nabla^2\mathbf{u}_s, \quad (71)$$

and a compressible (potential) part $\mathbf{u}_p = \nabla\varphi$, where φ satisfies the equation

$$-\rho\omega^2\varphi = (\lambda + 2\mu)\nabla^2\varphi. \quad (72)$$

The solution of Eqs. (71) and (72) in spherical geometry, presented in Ref. 17, in our notations takes the form

$$u_r = -\frac{2}{r^2} \left[\tilde{a}\mathbf{e}^{ikr} \left(1 - \frac{1}{ikr} \right) + \frac{\tilde{b}}{r} \mathbf{e}^{i\tilde{k}r} \left(1 - i\tilde{k}r - \frac{1}{2}\tilde{k}^2r^2 \right) \right] \cos\theta, \quad (73)$$

$$u_\theta = -\frac{1}{r^2} \left[\tilde{a}\mathbf{e}^{ikr} \left(1 - ikr - \frac{1}{ikr} \right) + \frac{\tilde{b}}{r} \mathbf{e}^{i\tilde{k}r} (1 - i\tilde{k}r) \right] \sin\theta, \quad (74)$$

where $\tilde{k} = \omega/c_l$, $c_l^2 = (\lambda + 2\mu)/\rho$ is a speed of acoustical waves.

Comparison of Eqs. (73) and (74) for displacement components in compressible media with Eqs. (54) and (55) for incompressible media shows that influence of compressibility reduces to the correction of Eqs. (54) and (55), and this correction is on the order of $\tilde{k}R$. So, the relative correction is $\tilde{k}/k = c_t/c_l \sim 10^{-2} - 10^{-3}$ for media-like tissues that can be considered as incompressible. Therefore, Eq. (63) with precision on the order of $10^{-2} - 10^{-3}$ can be used to describe the displacement of the solid sphere in soft tissues.

B. Transient displacement of a gas bubble

The difference between the solid sphere and bubble displacement analysis is the boundary conditions. The boundary condition for the static bubble displacement given by Eq. (32) is also valid for transient displacement. Equations (54) and (55) for displacement components are used to calculate $\sigma_{r\theta}$, which at the bubble surface $r=R$ is equal to

$$\sigma_{r\theta} = -\frac{\mu}{R^3} \sin\theta \left[\tilde{a}\mathbf{e}^{ikR} \left(3ikR + k^2R^2 - 6 + \frac{6}{ikR} \right) - \frac{6\tilde{b}}{R} \right] = 0. \quad (75)$$

In this case \tilde{b} is

$$\tilde{b} = \tilde{a}\mathbf{e}^{ikR} R \left(\frac{1}{2}ikR + \frac{1}{6}k^2R^2 - 1 + \frac{1}{ikR} \right), \quad (76)$$

and the bubble surface displacement components estimated with Eqs. (54) and (55) are

$$u_r = u_\omega \left(1 - \frac{1}{3}ikR \right) \cos\theta, \quad (77)$$

$$u_\theta = -\frac{u_\omega}{2} \left(1 + \frac{1}{3}ikR \right) \sin\theta. \quad (78)$$

The variable u_ω introduced here can be interpreted as low-frequency displacement

$$u_\omega = -\frac{ik}{R} \tilde{a}\mathbf{e}^{ikR}. \quad (78)$$

The radial component of the stress tensor σ_{rr} at the bubble surface is

$$\sigma_{rr} = -\frac{2\mu}{R} u_\omega (1 - ikR) \cos\theta. \quad (79)$$

Using Eq. (57) for the pressure and Eq. (59) for $\nabla^2 f + k^2 f$, and \tilde{b} defined in Eq. (76), the pressure in the medium around the moving bubble is

$$p = p_0 + \frac{\mu}{R} u_\omega \left(1 - ikR - \frac{1}{2}k^2R^2 + \frac{1}{6}ik^3R^3 \right) \cos\theta. \quad (80)$$

Equation (19) accounts for the force on the bubble oscillating in the elastic medium

$$F_\omega = -4\pi\mu R u_\omega \left(1 - ikR - \frac{1}{6}k^2R^2 + \frac{1}{18}ik^3R^3 \right). \quad (81)$$

When the external force $F_\omega^{(\text{ext})}$ (i.e., the radiation force) is applied to the bubble, Eq. (81) describes the bubble movement where $F_\omega^{(\text{ext})} = -F_\omega$

$$4\pi\mu R u_\omega (1 - ikR - \frac{1}{6}k^2 R^2 + \frac{1}{18}ik^3 R^3) = F_\omega^{(\text{ext})}. \quad (82)$$

Equation (82) describes bubble oscillations induced by the external force. In the time domain the equation for bubble oscillations looks like

$$U + \dot{U} + \frac{1}{6}\ddot{U} + \frac{1}{18}\dddot{U} = \tilde{F}_{\text{ext}}. \quad (83)$$

Equation (83) is written in dimensionless variables: U is the displacement normalized R , \tilde{F}_{ext} is the dimensionless external force, and dots above the function U denote the derivatives with respect to dimensionless time τ

$$U = \frac{u(t)}{R}, \quad \tilde{F}_{\text{ext}} = \frac{F_{\text{ext}}}{4\pi\mu R^2}, \quad \tau = \frac{tc_t}{R}. \quad (84)$$

Equations (82) and (83) derived for bubble displacement in incompressible media can be modified to take compressibility into account. As shown above, however, the correction is on the order of 10^{-2} – 10^{-3} . In addition, inclusion of compressibility makes the equation in the frequency domain very bulky and does not allow one to rewrite it in the time domain.

Equation (83) can be solved by combining a general solution of the homogeneous equation with a particular solution of the inhomogeneous equation. We take the general solution as

$$U = C_1 e^{\lambda_1 \tau} + C_2 e^{\lambda_2 \tau} + C_3 e^{\lambda_3 \tau}. \quad (85)$$

Here

$$\begin{aligned} \lambda_1 &= -1.133\,176, & \lambda_2 &= -0.933\,412 - 3.874\,701i, \\ \lambda_3 &= -0.933\,412 + 3.874\,701i, \end{aligned} \quad (86)$$

are roots of the algebraic equation

$$\lambda^3 + 3\lambda^2 + 18\lambda + 18 = 0, \quad (87)$$

and C_1 , C_2 , C_3 are constants. Since all three roots of the characteristic equation have negative real parts, the displacements are decreasing with time. The general solution, however, can be determined by variations of C_1 , C_2 , and C_3 .

The special case when the external force $\tilde{F}_{\text{ext}}(t)$ is a δ function is considered here. Therefore, C_1 , C_2 , and C_3 are the constants determined by initial conditions

$$\frac{1}{18}\ddot{U} = 1; \quad \dot{U} = 0; \quad U = 0. \quad (88)$$

The normalized solution of Eq. (83) is plotted in Fig. 3 (solid line). The external force in Eq. (83) is taken as a δ function. The bubble displacement has oscillatory character.

For comparison, the displacement of the weightless solid sphere (dashed line) is also shown in Fig. 3—there are no oscillations.

As discussed in the previous section, the bubble changes its shape during displacement. In process of transient displacement, the bubble shape is also changed. Equations (77) show that both components u_r and u_θ of bubble surface displacement as well as u depend on frequency since k is frequency dependent. Therefore, bubble deformation is more

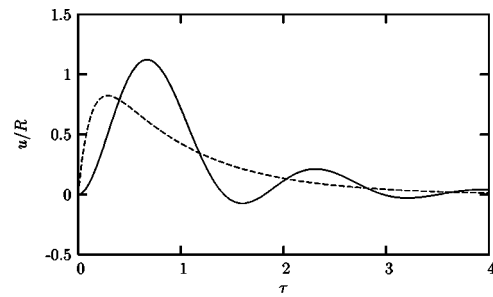


FIG. 3. Responses of the bubble (solid line) and the solid sphere (dashed line) to the short pulse force taken as a δ function, $\tau = tc_t/R$ here is dimensionless time.

complicated during transient displacement compared to the static case. For example, consider displacement of points on the bubble surface that correspond to $\theta = 0$. These points move back and forth along the r axis, which goes through the bubble center in direction of the force. Their displacement is characterized by \tilde{u}_r

$$\tilde{u}_r = (1 - \frac{1}{3}ikR)u_\omega. \quad (89)$$

The points on the bubble surface with $\theta = \pi/2$ also move, and their displacement is \tilde{u}_θ

$$\tilde{u}_\theta = (1 + \frac{1}{3}ikR)u_\omega. \quad (90)$$

Dependence of $u(t)$, as well as r and θ components of displacement in time representation $u_r^{(t)}$ and $u_\theta^{(t)}$ on dimensionless time in the elastic incompressible inviscous medium, are plotted in Fig. 4.

The response of the small bubble to the short pulsed force will be similar to the solid sphere response as the small bubble shape is stabilized by the surface tension. Additional investigation should be done to determine when the surface tension starts playing an important role.

IV. BUBBLE DISPLACEMENT IN VISCOELASTIC MEDIA

In order to determine the displacement of the solid sphere or the gas bubble in the real elastic media, the viscosity of the medium should be taken into account. This can be done by using the approach developed for the solid sphere in tissue.⁷ In our procedure the shear modulus μ should be replaced by $(\mu - i\omega\eta)$, where η is a shear viscosity coefficient. For displacements of the gas bubble in the viscoelastic medium, Eq. (82) takes the form

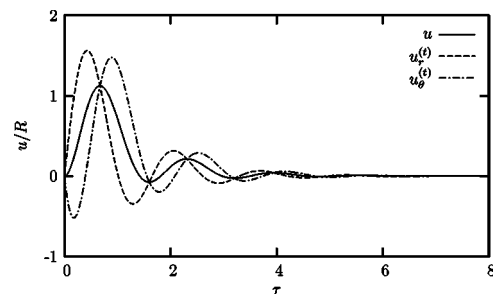


FIG. 4. Dependence of u , $u_r^{(t)}$ and $u_\theta^{(t)}$ on dimensionless time $\tau = tc_t/R$ in an inviscous medium.

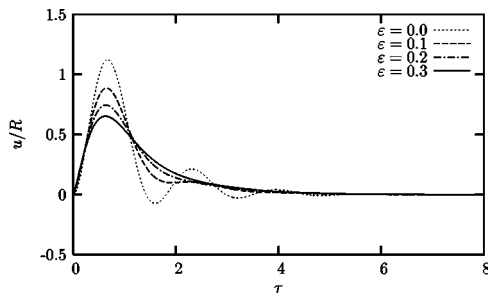


FIG. 5. Responses of the bubble to the short pulse force in the different viscoelastic media; the dimensionless parameter $\varepsilon = c_t \eta / R \mu$ takes value 0; 0.1; 0.2; 0.3.

$$4\pi(\mu - i\omega\eta)Ru_\omega(1 - ikR - \frac{1}{6}k^2R^2 + \frac{1}{18}ik^3R^3) = F_\omega^{(\text{ext})}, \quad (91)$$

where k is

$$k = \frac{\omega}{c_t \sqrt{1 - i\omega\eta/\mu}}. \quad (92)$$

Equation (91) allows for the estimation of the bubble response to the external force in the viscoelastic medium. Here, Eq. (91) is used to estimate bubble displacement when the external force $F_{\text{ext}}(t)$ as a δ function is applied. Since F_ω^{ext} is const for the δ function, it is easy to calculate u_ω using Eq. (91) and then apply inverse Fourier transform to get the bubble response to very short pulse. The bubble responses to the short pulsed force in different media are shown in Fig. 5. The calculations are performed in the frequency domain with inverse Fourier transformation. The number of harmonics in calculations has been equal to 1024. Normalization for t is the same as earlier and is given by Eq. (84). The dimensionless parameter ε is introduced to characterize the ratio between viscosity and elasticity

$$\varepsilon = \frac{c_t \eta}{\mu R}. \quad (93)$$

The results shown in Fig. 5 have an obvious interpretation: under the short impulse force the gas bubble displacement in the medium with no viscosity has oscillatory character. Viscosity eliminates oscillations in bubble displacement and reduces the maximum value of displacement.

Dependence of the variables u , $u_r^{(t)}$, and $u_\theta^{(t)}$ on dimensionless time in the viscoelastic medium with $\varepsilon=0.2$ is shown in Fig. 6. Comparison of the plots in Figs. 6 and 4

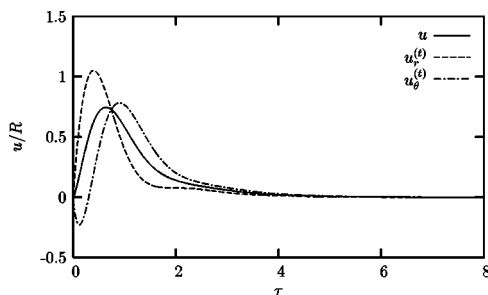


FIG. 6. Dependence of u , $u_r^{(t)}$, and $u_\theta^{(t)}$ on dimensionless time in the viscoelastic medium, the external force is a δ function and dimensionless parameter $\varepsilon=0.2$.

shows that there are no oscillations of u , $u_r^{(t)}$, and $u_\theta^{(t)}$ in the viscoelastic medium for up to $\varepsilon=0.2$, and the peak values of all variables in Fig. 6 are less than the ones in Fig. 4.

V. DISCUSSION AND CONCLUSIONS

Previously, the equilibrium shape of an acoustically levitated bubble in liquid has been investigated theoretically.¹⁸ In that case, the acoustic radiation pressure on the bubble surface counteracts gravity, and there is no dipole pressure and bubble deformation caused by the quadrupole term in pressure. In the problem considered here, the bubble shape changes due to the dipole term in pressure.

For transient displacement of solid sphere, our approach agrees with those presented earlier.⁷ Indeed, Eq. (63) couples the applied force and the displacement, while the relationship between pressure and velocity was derived for an oscillating sphere.⁷ The derived formula for impedance coincides with $F_\omega^{(\text{ext})}/u_\omega$ obtained from Eq. (63) if the elastic medium is incompressible, frictionless, and only induced mass is taken into account.

In this paper a general approach is developed to estimate the static displacement of a solid sphere and a gas bubble induced by an external force in elastic incompressible media. The method is similar to a Stokes' force derivation in viscous incompressible liquid. Displacement of the solid sphere is given by Eq. (31) and components of bubble displacement are determined by Eq. (43). Calculations show that the bubble is not only displaced by the applied force but also changes its shape. The initially spherical bubble has a shape described by Eq. (49) after displacement.

For transient displacement of a solid sphere as a bubble in elastic media induced by a time-dependent force, the motion equations are derived in the frequency domain as well as in the time domain. The motion equations for a rigid sphere in an inviscous elastic medium are Eq. (63) in the frequency domain and Eq. (68) in the time domain. They can be solved analytically or numerically for any external force applied to the solid sphere. Here, the solution is presented for the external force as a short pulse.

The transient bubble displacement in elastic media is described by Eq. (82) in the frequency domain and Eq. (83) in the time domain. Equation (83) has been solved analytically and presented in Fig. 3 showing oscillatory behavior in response to a short impulse force.

Since the effect of viscosity on a solid sphere is more or less known, only bubble displacement in the viscoelastic medium is considered here. Analysis is performed in the frequency domain by applying an inverse Fourier transform method. The result is viscosity eliminates oscillations in bubble response to a short impulse force and reduces the maximum value of bubble displacement.

In the future, nonlinear displacement of a solid sphere as well as a gas bubble in elastic media and influence of inhomogeneity of the elastic material will be studied.

ACKNOWLEDGMENTS

This work was supported by NIH Grant 1 R03 EB004047-01 and the Internal Research and Development Funds at the Applied Research Laboratories, The University of Texas at Austin.

- ¹A. P. Sarvazyan, O. V. Rudenko, S. D. Swanson, J. B. Fowlkes, and S. Y. Emelianov, "Shear wave elasticity imaging: A new ultrasonic technology of medical diagnostics," *Ultrasound Med. Biol.* **24**, 1419–1435 (1998).
- ²W. F. Walker, "Internal deformation of a uniform elastic solid by acoustic radiation force," *J. Acoust. Soc. Am.* **105**, 2508–2518 (1999).
- ³K. R. Nightingale, M. L. Palmeri, R. W. Nightingale, and G. E. Trahey, "On the feasibility of remote palpation using acoustic radiation force," *J. Acoust. Soc. Am.* **110**, 625–634 (2001).
- ⁴J. Bercoff, M. Tanter, and M. Fink, "Supersonic shear imaging: A new technique for soft tissue elasticity mapping," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **51**, 396–409 (2004).
- ⁵T. N. Erpelding, R. C. Booi, K. W. Hollman, and M. O'Donnell, "Measuring tissue elastic properties using acoustic radiation force on laser-generated bubble," *Proc. IEEE Ultrasonics Symposium*, 554–557 (2003).
- ⁶S. Chen, M. Fatemi, and J. F. Greenleaf, "Remote measurement of material properties from radiation force induced vibration of an embedded sphere," *J. Acoust. Soc. Am.* **112**, 884–889 (2002).
- ⁷H. L. Oestreicher, "Field and impedance of an oscillating sphere in a viscoelastic medium with an application to biophysics," *J. Acoust. Soc. Am.* **23**, 707–714 (1951).
- ⁸A. Prosperetti, "Bubble phenomena in sound fields. II," *Ultrasonics* **22**, 115–124 (1984).
- ⁹T. G. Leighton, *The Acoustic Bubble* (Academic, San Diego, 1994).
- ¹⁰P. J. Westervelt, "Acoustic radiation pressure," *J. Acoust. Soc. Am.* **29**, 26–29 (1957).
- ¹¹Yu. A. Ilinskii and E. A. Zabolotskaya, "Cooperative radiation and scattering of acoustic waves by gas bubbles in liquid," *J. Acoust. Soc. Am.* **92**, 2837–2841 (1992).
- ¹²T. Hasegawa and K. Yosioka, "Acoustic-radiation force on a solid elastic sphere," *J. Acoust. Soc. Am.* **46**, 1139–1143 (1969).
- ¹³Yu. A. Ilinskii, G. D. Meegan, and E. A. Zabolotskaya, "Bubble translation and deformation induced by ultrasound radiation force," *J. Acoust. Soc. Am.* **115**, 2560 (2004).
- ¹⁴L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*, 2nd ed. (Pergamon, New York, 1987).
- ¹⁵A. C. Eringen and E. S. Suhubi, *Elastodynamics* (Academic, New York, 1974), Vol. 1.
- ¹⁶L. D. Landau and E. M. Lifshitz, *Theory of Elasticity*, 3rd ed. (Pergamon, New York, 1986).
- ¹⁷H. Lamb, *Hydrodynamic*, 6th ed. (Dover, New York, 1944).
- ¹⁸T. J. Asaki and P. L. Marston, "Equilibrium shape of an acoustically levitated bubble driven above resonance," *J. Acoust. Soc. Am.* **97**, 2138–2143 (1995).

The mechanical and thermal effects of focused ultrasound in a model biological material

Feng Feng and Ajit Mal^{a)}

Mechanical and Aerospace Engineering Department, University of California, Los Angeles, California 90095

Michael Kabo

College of Engineering and Computer Science, California State University, Northridge, California 91330

Jeffrey C. Wang

Department of Orthopaedics and Neurosurgery, University of California, Los Angeles, California 90095

Yoseph Bar-Cohen

Jet Propulsion Laboratory/Caltech, Pasadena, California 91109

(Received 20 September 2004; revised 21 January 2005; accepted 28 January 2005)

This paper is motivated by possible medical applications of focused ultrasound in the minimally invasive treatment of a variety of musculoskeletal disorders that are responsive to thermal treatment. A model-based analysis of the interaction of high-intensity focused ultrasound with biological materials is carried out in an effort to predict the path of the sound waves and the temperature field in the focal region. A finite-element-based general purpose code called PZFlex is used to determine the effects of nonlinearity and geometrical complexity of biological structures. It was found that at frequencies of interest in therapeutic applications, the nonlinear effects are usually negligible and the geometrical complexities can be handled through a substructuring procedure. An approximate analytical method with acceptable accuracy is developed as an alternative to the purely numerical approach used in PZFlex. The mechanical and thermal effects in two-layered fluid material systems induced by high-frequency focused ultrasound are calculated through this analytical method. The results are compared with those obtained using PZFlex as a benchmark. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1873372]

PACS numbers: 43.80.Sh, 43.35Ac [FD]

Pages: 2347–2355

I. INTRODUCTION

The interaction between ultrasound and biological tissues has been the subject of numerous investigators for over three decades (Bilgen and Insana, 1996; Fry and Johnson, 1978; Hutchinson and Hynynen, 1996; Lele and Pierce, 1973; Rivens *et al.*, 1996; Sanghvi *et al.*, 1996; Seip *et al.*, 1996; Wall *et al.*, 1951; Zaneli *et al.* 1994). The application of focused ultrasound as a minimally invasive surgical tool has been demonstrated experimentally (Fry and Johnson, 1978), explained theoretically (Lele and Pierce, 1973; Robinson and Lele, 1972; Damianou, 1997), and practiced clinically (Hallaj *et al.*, 2001). Although the number and scope of the therapeutic applications of high-intensity focused ultrasound have grown in recent years, the technique is still far from being widely accepted by the clinical community. Since the target area is usually quite small, the focal spot generated by the ultrasound transducer needs to be small (Hynynen, 1991). Moreover, in order to avoid damage to neighboring healthy tissues, an effective control of the path of the ultrasound beam and the identification of the location of the focal spot is essential in clinical applications. Thus further progress in the noninvasive or minimally invasive use of ultrasound for clinical applications will require advances in a number of enabling technologies, such as phased array fab-

rication and other techniques of controlling the beam shape, as well as a better understanding of the interaction of focused ultrasound with various types of tissues.

The thermal and mechanical interaction of focused ultrasound with biological tissues and structures is considered in this paper. Since experimental studies using living tissue are difficult and costly, theoretical simulation of the problem can be extremely useful in providing a firm scientific basis for future clinical investigations of focused ultrasound. In previous studies (Mal and Feng, 2002, 2003), relatively complex geometrical and acoustic properties of human tissues have been considered by using the finite-element-based software called PZFlex (2001) in an effort to obtain quantitative information on the interaction between high-energy ultrasound and biological structures. A result of these studies is that the nonlinear effects in the focal region are negligible in the frequency and pressure ranges of interest. However, a number of difficulties were encountered in this purely numerical approach. A major difficulty encountered was the extremely high storage requirement and runtime needed to compute the ultrasound field in the required frequency range and geometrical properties of the model.

Several approximate analytical solutions of the problem of acoustic wave radiation from sources in a homogeneous ideal fluid are available in the literature (King, 1934; O'Neil, 1949; Williams and Labaw, 1945; Williams, 1946, 1947). Williams and O'Neil evaluated the radiated pressure field by computing the Rayleigh integral on the curved surface of the

^{a)} Author to whom correspondence should be addressed. Electronic mail: ajit@seas.ucla.edu

radiator. For a planar source the Rayleigh integral is the exact formulation of the Huygens principle (Morse and Feshbach, 1953). Due to the planar nature of the surface, the spherical waves emitted by all these secondary sources do not interact with the surface itself. However, if the surface is curved, then the spherical waves are diffracted by the surface. Consequently, the amplitude of one spherical wave at one source point depends not only on the initial wave loaded on the local surface, but also on all other diffracted waves reaching that point.

Another interesting analytical method was presented by Killer (1989) and Coulouvrat (1993). It is based on two expansions of the pressure field in spherical harmonics in conveniently chosen domains that are connected by an artificial boundary. This method is “exact” in the sense that it explicitly considers the actual boundary conditions on the real curved surface; approximations only arise from the numerical discretization.

In this paper, an approximate analytical method is developed as an alternative to the purely numerical approach. The method is applicable to cases where the wavelength of the ultrasound is small compared to the geometrical features of the biological structure of interest. The method is based on the approach presented by Killer (1989) and Coulouvrat (1993), and can be considered to be a generalization of their methods to damped multilayered systems. The approach is used to determine the mechanical and thermal fields in a homogenous viscous fluid and in a two-layered material system with planar interfaces. The method can be extended to deal with curved interfaces and multilayered media in a straightforward manner. The developed method is sought to support the increasing application of focused ultrasound to perform minimally invasive medical treatment of musculoskeletal disorders (Bar-Cohen *et al.*, 1999), as well as cancer (Ter Haar *et al.*, 1989).

II. THE APPROXIMATE ANALYTICAL METHOD

A conceptual focused ultrasound system is sketched in Fig. 1 where a spherical cap transducer is used as the source located in a sealed chamber containing a coupling fluid. A thin membrane at the bottom of the chamber separates the fluid from the biologic tissue. An axisymmetric analog problem shown in Fig. 2 is extracted from the therapeutic problem of Fig. 1 for theoretical treatment. Here P is the observation point, and Γ_S is the spherical cap transducer with transverse diameter, $2a$, focal length, f , and aperture angle, β . A stiff baffle, Γ_B , is assumed to be located on the horizontal plane outside the lens. The fluid (e.g., water) occupies the domain, $\Omega = \Omega_i \cup \Omega_o$, and a biological material (e.g., muscle) is in the domain Ω_H with a plane interface, Γ_I , separating it from the fluid. The top of the spherical cap, O , is chosen as the origin of a cylindrical coordinate system, (R, φ, z) . The point O' is the geometrical focus of the lens (i.e., $f = OO'$), and h is the height of the lens. The point C is at the intersection between the baffle plane and the acoustic axis Oz . The inner and outer domains, Ω_i and Ω_o , are separated by an imaginary hemisphere, Γ_A , of the radius, a , centered at the origin C . A spherical coordinate system, (r, θ, ϕ) is introduced with the origin at the point C .

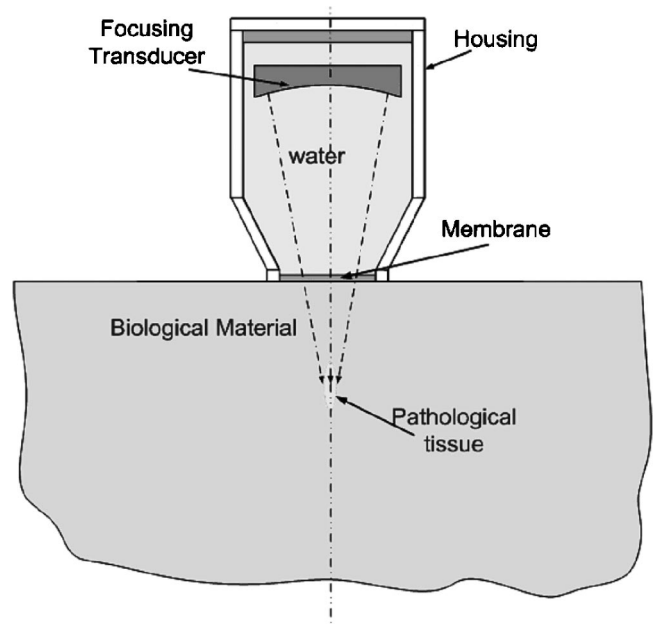


FIG. 1. A conceptual focused ultrasound system for therapeutic application. Note: The wave path shown in this figure is a straight line with no refraction due to the assumption that the wave speed in the tissues is the same as in water.

The solution of the two-layered material problem is realized in two steps. The first step is to calculate the pressure and temperature fields in a fluid occupying the domain $(\Omega_i \cup \Omega_o \cup \Omega_H)$ mathematically but with the original boundary conditions of the model in Fig. 2. In the second step the pressure, velocity, and temperature fields in the biological material in the two-layered model of Fig. 2 are evaluated by applying the boundary conditions at the interface, Γ_I , which are derived from the first step with suitable modifications to account for reflection and transmission at the interface.

A. Homogeneous material

The fluid in the domain $\Omega_i \cup \Omega_o \cup \Omega_H$ is assumed to be homogeneous, isotropic, and viscous. As indicated earlier,

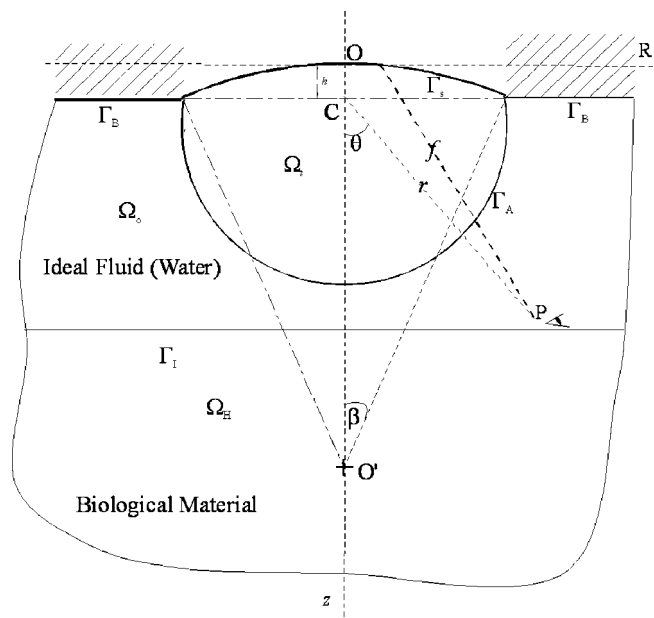


FIG. 2. Geometry of the bimaterial problem.

TABLE I. Mechanical and thermal properties in material system I.

Material	ρ (g/cc)	c_p (mm/ms)	c_s (mm/ms)	a (dB/cm)	b	C_p (J/kg K)	k_h (W/m K)
Water	1000	1500	0	0	2.0	4200	0.588
Muscle	1041	1571	0	0.55 ^a	1.0	3720	0.54

^aAt 1 MHz.

nonlinear effects can be neglected in most applications considered here so that the time harmonic pressure field, p , in the fluid satisfies the Helmholtz equation,

$$\nabla^2 p + k^2 p = 0 \quad (1)$$

in Ω where $k = \omega/c$, ω is the circular frequency, and k is the wave number. Here as in all subsequent equations the time-dependency factor $e^{i\omega t}$ is omitted.

The general solution for the pressure field in spherical coordinates can be expressed in terms of the spherical Hankel functions $h_n^{(1)}(kr)$ and $h_n^{(2)}(kr)$ (Morse and Feshbach, 1953) in the outer domain $\Omega_o \cup \Omega_H$ as

$$p^{(o)} = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \{ [a_{1m} \sin(m\phi) + a_{2m} \cos(m\phi)] P_n^m(\cos \theta) [b_{1n} h_n^{(1)}(kr) + b_{2n} h_n^{(2)}(kr)] \} \quad (2)$$

and in terms of the spherical Bessel functions $j_n(kr)$ and $n_n(kr)$ in the inner domain Ω_I as

$$p^{(i)} = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \{ [a_{1m} \sin(m\phi) + a_{2m} \cos(m\phi)] P_n^m(\cos \theta) [b_{1n} j_n(kr) + b_{2n} n_n(kr)] \}. \quad (3)$$

In Eqs. (2) and (3), P_n is the n th-order Legendre polynomial, and a_{1m} , a_{2m} , b_{1n} , and b_{2n} are unknown constants to be determined from the boundary conditions.

Taking into account the axial symmetry of the problem, the boundary condition on the baffle plane Γ_B , the boundedness condition in the interior domain, and the radiation condition in the outer domain, the pressures $p^{(i)}$ and $p^{(o)}$ can be simplified as

$$p^{(i)} = \sum_{n=0}^{\infty} \alpha_n^{(i)} j_n(kr) P_n(\cos \theta) \quad \text{in } \Omega_i, \quad (4)$$

$$p^{(o)} = \sum_{n=0}^{\infty} \alpha_n^{(o)} h_n^{(2)}(kr) P_n(\cos \theta) \quad \text{in } \Omega_o \cup \Omega_H, \quad (5)$$

where $\alpha_n^{(i)}$ and $\alpha_n^{(o)}$ are unknown inner and outer coefficients. Applying the continuity conditions on the artificial boundary at a finite number of collocation points (Coulouvrat, 1993), the pressure field in a homogenous fluid material can be obtained through the solution of the linear system

$$\mathbf{S}\mathbf{x} = \mathbf{b} \quad (6)$$

in which $[\mathbf{S}]$ is a known matrix, $\{\mathbf{x}\}$ is a column vector containing the unknown inner coefficients, ($x_n = \alpha_n^{(i)}$), and $\{\mathbf{b}\}$ is

the column vector describing the known applied pressure field at the collocation points. It should be noted that in general the vectors $\{\mathbf{b}\}$ and $\{\mathbf{x}\}$ are complex and the matrix $[\mathbf{S}]$ is, in general, full, nonsymmetric, and complex. The dimension of the system (6) depends on the number of collocation points.

Generally, the phase velocity or the wave number of a material can be assumed to be a complex number for time harmonic wave propagation problems involving materials with dissipation. As an example, the particle motion, $A(x)$, associated with a plane wave propagating along the positive x -direction in a dissipative material can be obtained by simply replacing the wave number k in the motion equation in a perfect fluid with a complex number $k_r + ik_i$:

$$A(x) = A_0 e^{-ikx} = A_0 e^{-ik_r x} \cdot e^{-\alpha x}, \quad (7)$$

where A_0 is the amplitude, $k = k_r - i\alpha$, and α is the attenuation coefficient, which must be positive to represent the exponential decay in the amplitude of the wave. The attenuation of the waves is often described by decibels (dB) in acoustics:

$$\text{dB} = 20 \log_{10} \left(\frac{A_0}{A(x)} \right). \quad (8)$$

Thus

$$\alpha = \frac{\text{dB}}{20 \log_{10} e} = \frac{\text{dB}}{8.686}. \quad (9)$$

The solution of any time harmonic plane wave propagation problem in a dissipative medium can be obtained from that for a nondissipative medium by simply replacing k by $k - i\alpha$. As an example, to calculate the wave field in a muscle material at 100 kHz, the related material properties are given in Table I and the frequency dependence of α is given in the following equation (Duck, 1990).

$$\alpha = a f^b, \quad (10)$$

where a and b are material constants and f is normalized frequency with the center frequency (Duck, 1990). Thus $\alpha = 0.576$ and $k = 399.95 - 0.576i$. The pressure field in a dissipative medium can be obtained by substituting this value of k in Eq. (6), solving for the unknown coefficients, and evaluating the expressions (4) and (5).

One purpose of focused ultrasound in the clinical applications is to deliver controlled heat to the pathological tissue, and the shrinkage of the pathological tissue caused by the controlled heat can, in appropriate cases, restore the tissue to its original state. The most important tasks in the modeling

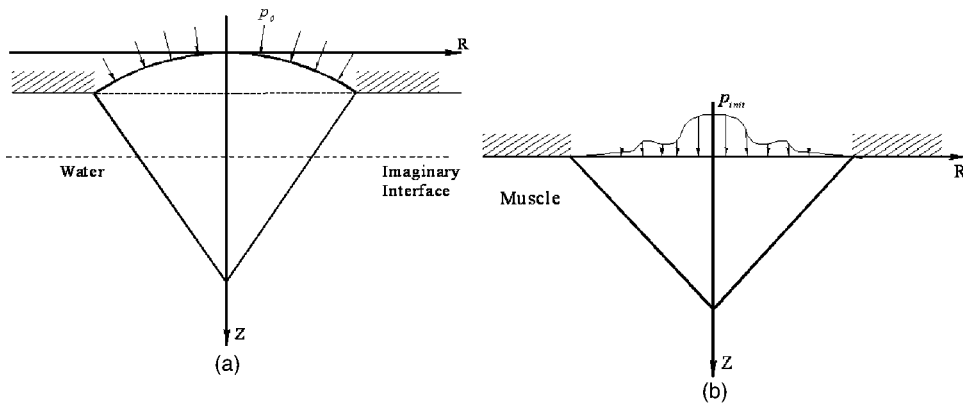


FIG. 3. Simplified subproblems.

study are to predict the location and extent of the focal zone, the temperature distribution, and, in particular, the highest temperature in the focal zone.

The average energy dissipation, dE , in a cycle per unit length of a plane wave (O'Neil, 1949) can be expressed in the form

$$dE = \text{Re} \left[\frac{p\bar{v}}{2} \right] dx dy, \quad (11)$$

where p is the pressure, \bar{v} is the complex conjugate of the velocity, v , and "Re" implies the real part of the expression. The average energy loss, E_{loss} , in a cycle per unit volume is

$$E_{\text{loss}} = \alpha \text{Re}[p\bar{v}]. \quad (12)$$

Then the modified heat transfer equation is

$$\rho C_p \frac{\partial T}{\partial t} = \nabla \cdot (k_h \nabla T) + q, \quad (13)$$

where ρ is the density of the material, C_p is its specific heat, T is the temperature, t is the current time, k_h is the heat conductivity, and q is the acoustic energy dissipation per volume in unit time, which can be identified as E_{loss} defined in (12). Since the temperature field is only considered in a short time, the heat conduction effect is small enough to be ignored. Thus the temperature rise associated with the energy loss can be approximated as

$$T = \frac{\alpha \text{Re}[p\bar{v}]}{\rho C_p} t. \quad (14)$$

The above equation is the basis for calculating the thermal effects in a viscous fluid. It has been verified through numerical calculations using PZFlex that heat conduction results in small errors in the temperature field in the pressure and temperature ranges of interest here. The mechanical and thermal fields calculated by the present approximate approach will be validated by comparing them with the corresponding results using PZFlex.

B. Two-layered material model

The two-layered (water/muscle) model shown in Fig. 2 is considered next. The original model is separated into two submodels as sketched in Fig. 3. The first submodel shown in Fig. 3(a) is identical with the single material model, in this case water, discussed in the previous section. Dissipation in

water is ignored in the calculations. The imaginary surface shown as a dashed line in Fig. 3(a) is at the same location as the real interface, Γ_I , between two materials (Fig. 2). The pressure, p_o , and velocity, v_{in} , at the imaginary interface can be determined using the method presented in the previous section where p_o is the same as the pressure in the incident wave from water to the real interface in the two-layered model. The corresponding interfacial pressure, p_i , at the real interface Γ_I is determined by an approximate theory.

The calculation of the interfacial pressure p_i is based on the assumption that an infinitesimal region is locally subjected to an incident plane wave (Fig. 4). The reflection and transmission pressures and velocities at each collocation point on the interface Γ_I are evaluated through plane wave theory, after the local incident pressure and velocity at each collocation point are obtained from the first submodel. The interface conditions for the plane wave theory are

$$p_i(R, 0) = (1 + R_w) p_o(R, z_0), \quad (15)$$

$$p_i(R, 0) = T p_o(R, z_0), \quad (16)$$

in which T and R_w are transmission and reflection coefficients given by (see, e.g., Mal and Singh, 1991)

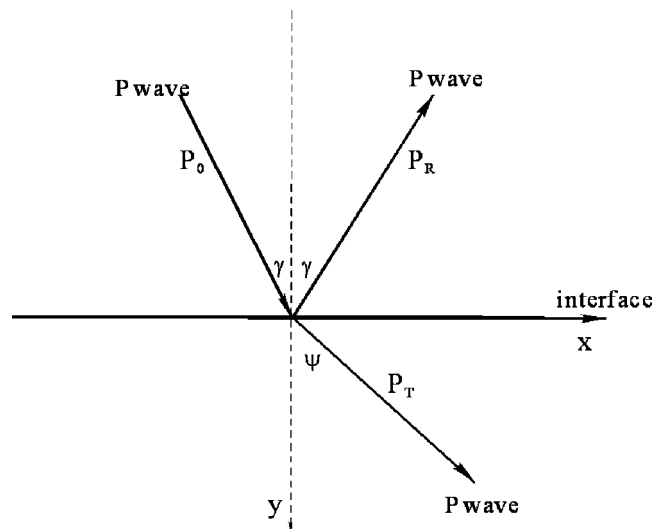


FIG. 4. Plane wave at the interface.

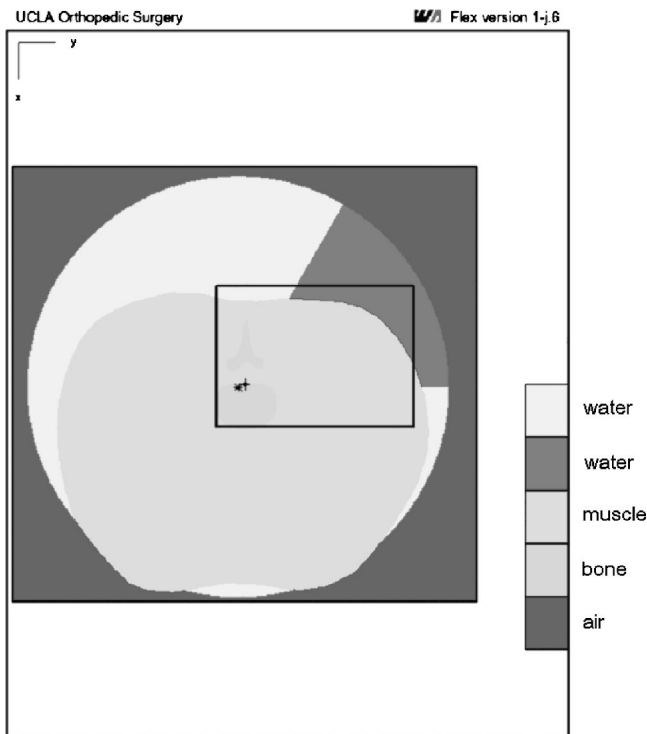


FIG. 5. FEM model for the therapeutic application.

$$T = \frac{2\rho_2 c_2 \cos \gamma}{\rho_2 c_2 \cos \gamma + \rho_1 c_1 \cos \psi},$$

$$R_w = \frac{\rho_2 c_2 \cos \gamma - \rho_1 c_1 \cos \psi}{\rho_2 c_2 \cos \gamma + \rho_1 c_1 \cos \psi},$$
(17)

where γ and ψ are incident and transmitted angles, ρ_i and c_i ($i=1,2$) are the densities and phase velocities of waves in the upper and lower medium, respectively.

The interfacial pressure, p_i , given by (15) or (16) is applied as the boundary condition in the second submodel shown in Fig. 3(b), which involves only one material (muscle) with damping. Then, the same procedure as for the single material model is used to characterize the mechanical and thermal fields of the second submodel. The pressure field is determined from the solution of Eq. (6) and then the thermal field is calculated approximately by Eq. (14).

III. ANALYTICAL RESULTS AND COMPARISON WITH THE FEM MODEL

A. Mechanical and thermal effects in a homogenous fluid

1. Mechanical effect and energy loss

The purpose of this study is to develop an analytical method to efficiently and accurately predict the sound field radiated by an acoustic lens (spherical cap), and to compare the results of computation with the FEM model by PZFlex (Fig. 5). Figures 6(a) and (b) show the comparison of the axial pressure and velocity (V_z) amplitudes assuming that the material is muscle with properties given in Table I, for a focal length of $f=0.3$ m and frequency of 100 kHz. It can be seen that the two results are indistinguishable in the plots. These results are confirmed by the results on the plane z

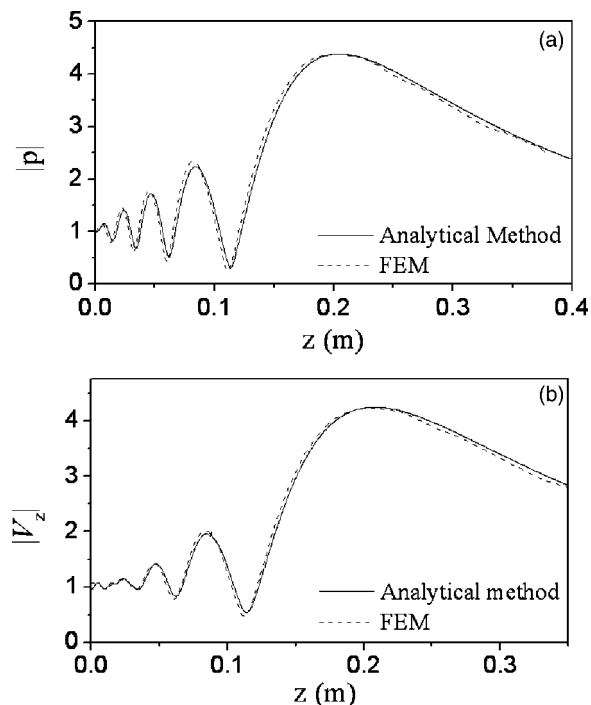


FIG. 6. (a) Pressure and (b) velocity (V_z) amplitudes along the axis, obtained by FEM and analytical method.

$=0.15$ m, shown in Figs. 7(a) and (b), showing excellent agreement, except at points very close to the source. The calculated energy losses in a cycle ($10 \mu s$) along the axis on $z=0.15$ m at the frequency 100 kHz are compared in Figs. 8(a) and (b). Again, the agreement is excellent except at field points near the lens. The reason for the discrepancies near the source is not completely clear, but they are probably caused by an insufficient number of elements in the FEM model.

These above results indicate that the analytical method

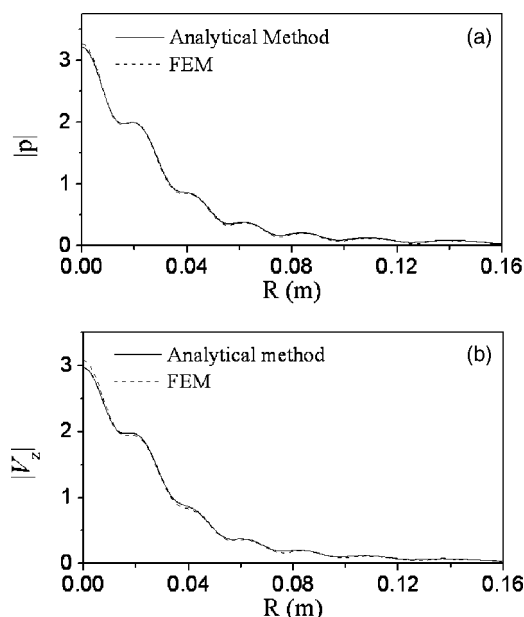


FIG. 7. (a) Pressure and (b) velocity (V_z) amplitude along the plane $z = 0.15$ m, obtained by FEM and analytical method.

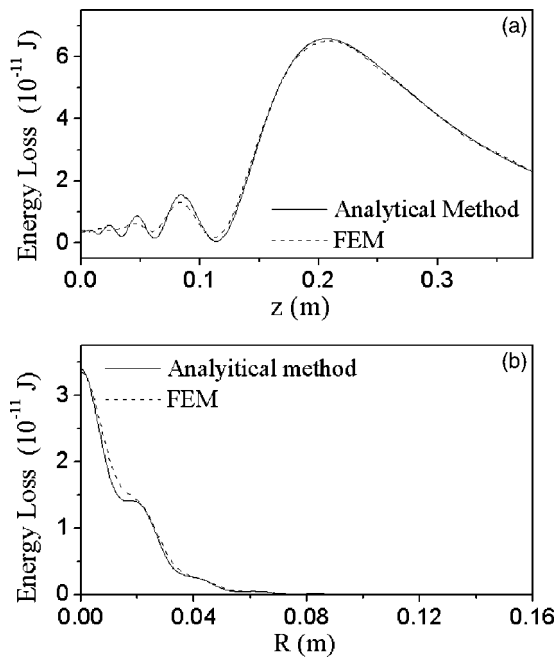


FIG. 8. Energy loss in one cycle (10 ms) along (a) the axis and (b) R direction at $z=0.15$ m.

can provide excellent results in a very short computing time, generally 2–5 s, as compared to FEM calculations that often take several hours (12 or more) to accomplish the same task. The mechanical effects in a dissipative material, such as muscle, can be calculated efficiently and accurately by the analytical method.

2. Analysis of focal zone

The dependence of the focal region on the frequency of the incident wave and the radius of the lens is determined by

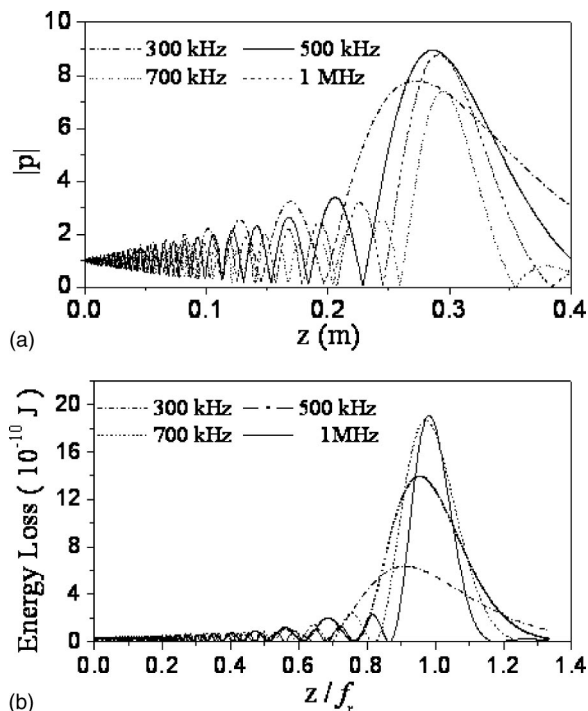


FIG. 9. Focus of pressure and energy loss at different frequencies.

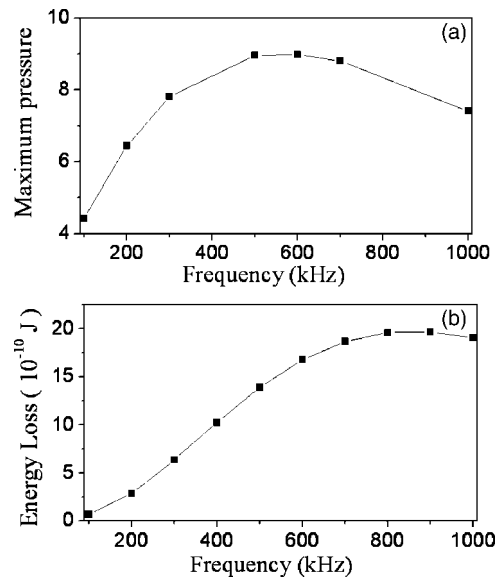


FIG. 10. Maximum pressure and energy loss at different frequencies.

the approximate analytical method for the single material (muscle) model. The pressure and energy loss (or temperature rise) along the lens axis at different frequencies are plotted in Figs. 9(a) and (b), respectively. They show that the focal region, where the pressure and temperature are the highest, becomes more localized (i.e., narrower) at higher frequencies. The maximum pressure and maximum energy loss in the focal region are plotted as functions of frequency in Figs. 10(a) and (b), respectively. The results indicate that the highest pressure and temperature occur at driving frequencies in the range of 700–1000 kHz.

The pressure field generated by a spherical lens in a material with damping exhibits different characteristics from those in a nondissipative material in which the pressure in the focal zone increases monotonously with frequency. Since the attenuation coefficient α is generally a function of frequency, the amplitude of the particle motion subjected to a plane wave decreases with increasing ultrasound frequency. But for a converging spherical wave in a dissipative material, the amplitude increases with propagation distance and increasing frequency due to the geometrical effect of focusing (nondamping effects) while it decreases with increasing distance and frequency due to energy loss in the material

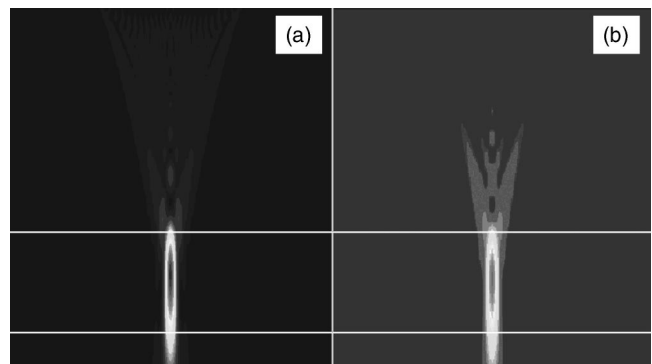


FIG. 11. Temperature fields at frequency 700 kHz by (a) analytical method and (b) FEM.

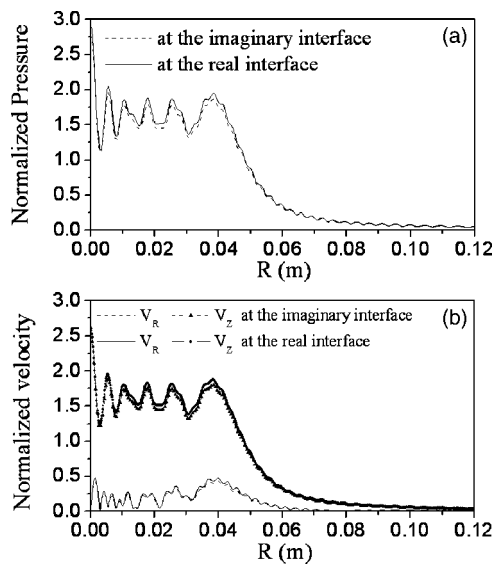


FIG. 12. (a) Pressure and (b) velocity at the interface between water and muscle (system I).

(damping effects). Thus, in a dissipative material the maximum pressure occurs at different spatial locations at different frequencies, as shown in Fig. 10(a). At high frequencies, the amplitude of pressure becomes negligibly small within a short distance away from the source (Mal *et al.*, 2003).

3. Temperature field

In order to calculate the temperature field, the uniform pressure of 0.25 MPa is assumed to be applied on the surface of the lens without loss of generality. The focal length of the lens is assumed to be 0.15 m. Typical results for the temperature fields obtained by the analytical method and FEM with an incident wave of frequency 700 kHz are compared at the end of 5-s exposure in Fig. 11.

The spatial temperature profiles are very similar for both solutions. The highest temperatures predicted in the muscle are 50.1 °C by the analytical method and 48.3 °C by the FEM. However, if a smaller value of the thermal conductivity (e.g., $k_h = 10^{-5}$ W/m K) is used for the material in the FEM calculation, a highest temperature of 49.2 °C is obtained. The difference between these two solutions is caused by the fact that heat conduction is ignored in the analytical method. The results are close enough for therapeutic applications. However, the computational needs are very different for these two methods—PZFlex requires over 10 h, while the analytical method needs only a few seconds to complete the calculations. Considering the tradeoff, the analytical method is clearly more useful for real-time clinical applications.

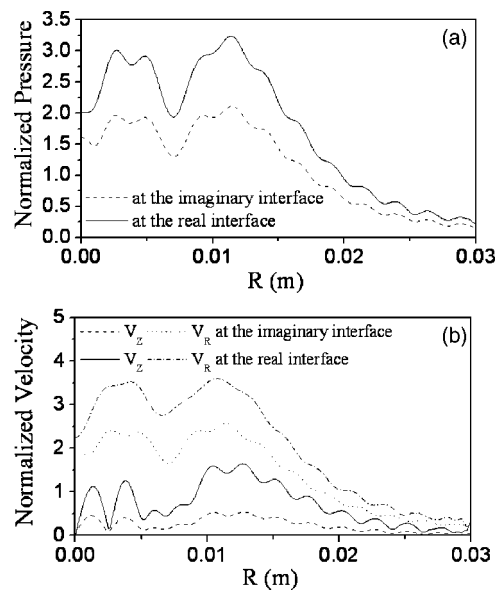


FIG. 13. (a) Pressure and (b) velocity at the interface between water and the imaginary material (system II).

B. Mechanical and thermal effects in a two-layered fluid material system

1. Interfacial pressure

As mentioned previously, in the two-step method, the interfacial pressure and velocity must be determined after solving the first subproblem. In this section, the results for the calculated pressure and velocity at the interface between water and muscle are plotted in Figs. 12(a) and (b). In this calculation, the focal length of the lens is 0.15 m and the depth of water is 0.075 m (the distance between the aperture plane of the lens and the interface). The driving frequency is 500 kHz and the (uniform) pressure on the surface of the lens is 0.25 MPa. It should be noted that since the wave impedances of water and muscle are very close, the plane wave transmission coefficient is close to 1, and the pair of curves in Fig. 12 are almost identical. In order to demonstrate the accuracy of the present method for two materials with large difference in their wave impedance, another pair (water and an imaginary material) described in Table II is used. The pressure and velocity distributions at the interface between water and the imaginary material are given in Figs. 13(a) and (b), in which the two curves are quite different due to the large mismatch in their wave impedances.

After the interfacial pressure and velocity have been obtained, we can now solve the second subproblem to determine the mechanical and thermal effects in the lower medium.

TABLE II. Mechanical and thermal properties in material system II.

Material	ρ (g/cc)	c_p (mm/ms)	c_s (mm/ms)	a (dB/cm)	b	C_p (J/kg K)	k_h (W/m K)
Upper (water)	1000	1500	0	0	2.0	4200	0.588
Lower ^a	1500	3000	0	40 ^b	1.0	3720	0.54

^aAn imaginary material.

^bAt 0.75 MHz.

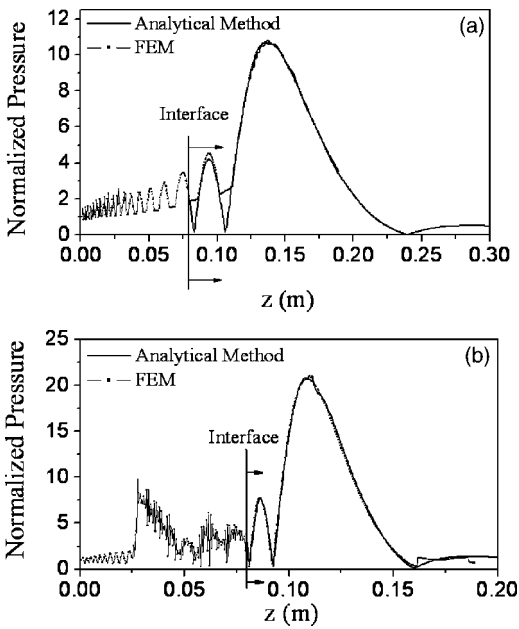


FIG. 14. The pressure distributions along the z direction in (a) system I and (b) system II obtained by FEM and the analytical method.

2. Two-layered material systems

In the next simulation, the mechanical and thermal effects in two material systems are calculated using the two-step method. One system (I) involves water and muscle, and the other (II) involves those listed in Table II. The focal length of the lens is 0.15 m, the water depth is 0.075 m, the frequency is 750 kHz, and the uniform pressure on the surface of the lens is 0.25 MPa.

In the previous section, the interfacial pressure and velocity have been determined. We can use our program to evaluate the mechanical and thermal effects in the second layer, representing the biomaterial. The validity of the approximate analytical method is demonstrated by comparison with the pressure along the z direction with that obtained from the FEM analysis (Fig. 14). The temperature fields in the two material systems will be compared in a contour plot with those obtained by the analytical method.

Figures 15 and 16 show the temperature fields in the lower material in these two systems. Again, the temperature distributions obtained by the two methods are very similar to each other. The maximum temperatures at the focus are I (55.3 °C and 53.7 °C) and II (51.7 °C and 53.0 °C) calculated from the approximate analytical method and FEM, respectively. However, the runtime for PZFlex is around 12 h or

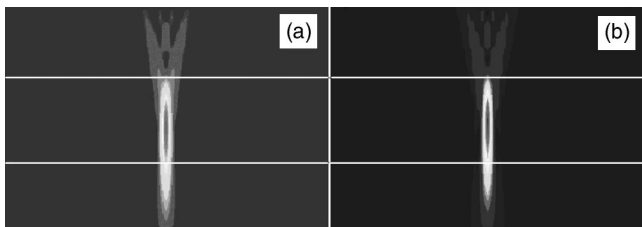


FIG. 15. The temperature fields in muscle (system I) obtained by (a) FEM and (b) analytical solutions.

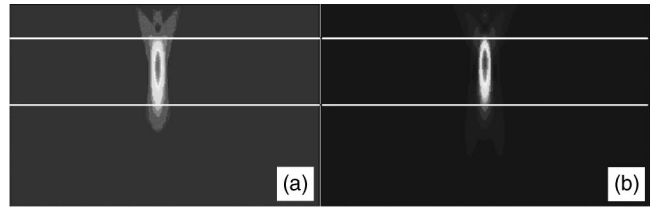


FIG. 16. The temperature fields in the imaginary material (system II) obtained by (a) FEM and (b) analytical solution.

more, but the analytical method needs only 3–5 s on the same computer.

The two black lines in Figs. 15 and 16 represent the top and bottom borders of the efficient heating region in the z direction. The profiles of the heated regions simulated by both methods are very similar to each other. Since heat transfer is considered in the FEM model, the width of the heating region is somewhat larger than that obtained from the analytical method.

IV. CONCLUDING REMARKS

A model-based analysis of the interaction of high-intensity focused ultrasound with biological materials was carried out in an effort to predict the path of the sound waves and the temperature field in the focal region. A novel approximate analytical approach has been developed as an alternative to the FEM code in an effort to improve the computational efficiency and accuracy of the simulations. The major features of the approach are (1) extension of the artificial boundary method developed by Coulouvrat (1993) to calculate mechanical and thermal effects in one homogeneous fluid with damping and (2) development of a two-step method using a local plane wave approximation to evaluate the mechanical and thermal effects in two-layered material systems. This analytical method works for any combination of fluidlike material and it can be extended to multilayered fluid and solid media with planar or curved interfaces.

The results presented here show that the approximate analytical method can be used to calculate the mechanical and thermal effects of focused ultrasound in two-layered material systems in real time. The calculated energy loss along the axis of the lens (z axis) shows that the input frequency not only determines the size and the temperature in the focal zone, but also the position of the focus which must coincide with the location of the pathological tissue in clinical applications. The analytical method can effectively and accurately evaluate the mechanical and thermal effects in the second layer in which the pathological tissues are likely to be located. In addition, the code based on the approximate analytical method is much more efficient than the FEM and provides results that are accurate to within 4%–5%, an acceptable error in clinical applications.

ACKNOWLEDGMENTS

This research was supported by a grant from the Whitaker Foundation.

Bar-Cohen, Y., Kabo, M., Kwong-Fu, H., Lee, P., Wang, J., and Mal, A.

- (1999). "Transcutaneous Spine Trauma and Disorders Treatment Using Ultrasonically Induced Confined Heat (ULICH) Zone," NASA New Technology Report, Item No. 0227b, Docket 20632, Patent disclosure Application No. 10/093866 (7 March, 2002).
- Bilgen, M., and Insana, M. F. (1996). "Effects of phase aberration on tissue heat generation and temperature elevation using therapeutic ultrasound," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **43**, 999–1010.
- Coulouvrate, F. (1993). "Continuous field radiated by a geometrically focused transducer: Numerical investigation and comparison with an approximate model," *J. Acoust. Soc. Am.* **94**, 1663–1675.
- Damianou, C. A., Sanghvi, N. T., and Fry, F. J. (1997). "Dependence of ultrasonic attenuation and absorption in dog soft tissues on temperature and thermal dose," *J. Acoust. Soc. Am.* **102**, 628–634.
- Duck, F. A. (1990). *Physical Properties of Tissue, a Comprehensive Reference Book* (Academic, San Diego).
- Fry, F., and Johnson, L. (1978). "Tumour irradiation with intense ultrasound," *Ultrasound Med. Biol.* **4**, 337–411.
- Hallaj, I. M., Cleveland, R. O., and Hynynen, K. (2001). "Simulations of the thermo-acoustic lens effect during focused ultrasound surgery," *J. Acoust. Soc. Am.* **109**, 2245–2253.
- Hutchinson, E. B., and Hynynen, K. (1996). "Intracavitary ultrasound phased arrays for noninvasive prostate surgery," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **43**, 1032–1042.
- Hynynen, K. (1991). "The threshold for thermally significant cavitations in dog's thigh muscle in vivo," *Ultrasound Med. Biol.* **17**, 157–169.
- Killer, J. B., and Givoli, D. (1989). "Exact non-reflecting boundary conditions," *J. Comput. Phys.* **82**, 172–192.
- King, L. V. (1934). "On the acoustic radiator field of the piezoelectric oscillator and the effect of the viscosity on transmission," *Can. J. Res.* **11**, 135–155.
- Lele, P. P., and Pierce, A. D. (1973). "The thermal hypothesis of the mechanism of ultrasonic focal destruction in organized tissues," *Interaction of ultrasound and biological tissues* FDA 73-8008 BRH/DBE (Bureau of Radiological Health, Washington, DC), pp. 121–128.
- Mal, A. K., and Singh, S. J. (1991). "Deformation of Elastic Solids" (Prentice Hall, New Jersey).
- Mal, A. K., Feng, F., Kabo, M., and Wang, J. (2002). "Interaction of focused ultrasound with biological materials," *Proc. SPIE, Smart Nondestructive Evaluation for Health Monitoring of Structural and Biological Systems*, Tribikram Kundu, Ed. 4702, pp. 339–348.
- Mal, A. K., Feng, F., Kabo, M., and Wang, J. (2003). "Pressure and Temperature Distribution in Biological Tissues by Focused Ultrasound," *Proc. SPIE, Smart Nondestructive Evaluation for Health Monitoring of Structural and Biological Systems*, Tribikram Kundu, Ed. 5047, pp. 339–348.
- Morse, P. M., and Feshbach, H. (1953). *Methods of Theoretical Physics* (McGraw-Hill, New York).
- O'Neil, H. T. (1949). "Theory of focusing radiators," *J. Acoust. Soc. Am.* **21**, 516–526.
- PZflex Software, Version: 1-j.6. Weidlinger Associates. Inc, 2001.
- Rivens, I. H., Clarke, R. L., and TerHaar, G. R. (1996). "Design of focused ultrasound surgery transducers," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **43**, 1023–1031.
- Robinson, T. C., and Lele, P. P. (1972). "An analysis of lesion development in the brain and plastics by high-intensity focused ultrasound at low-megahertz frequencies," *J. Acoust. Soc. Am.* **5**, 1333–1351.
- Sanghvi, N. T., Fry, F. J., Bihle, R., Foster, R. S., Phillips, M. H., Syrus, J., Zaitsev, A. V., and Hennige, C. W. (1996). "Noninvasive surgery of prostate tissue by high-intensity focused ultrasound," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **43**, 1099–1110.
- Seip, R., VanBaren, P., Cain, C. A., and Ebbini, E. S. (1996). "Noninvasive real-time multipoint temperature control for ultrasound phased array treatments," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **43**, 1063–1073.
- Ter Haar, G., Sinnett, D., and Rivens, I. (1989). "High intensity focused ultrasound—a surgical technique for the treatment of discrete liver tumours," *Phys. Med. Biol.* **34**, 1743–1750.
- Wall, P. D., Fry, W. J., Stephens, R., Tucker, D., and Lettvin, J. Y. (1951). "Changes produced in the central nervous system by ultrasound," *Science* **114**, 686–687.
- Williams, A. O. (1946). "Acoustic intensity distribution from a 'piston' source. II. The concave piston," *J. Acoust. Soc. Am.* **17**, 219–227.
- Williams, A. O. (1947). "Acoustic wave fronts from a 'piston' source," *J. Acoust. Soc. Am.* **19**, 156–161.
- Williams, A. O., and Labaw, L. W. (1945). "Acoustic intensity distribution from a 'piston' source," *J. Acoust. Soc. Am.* **16**, 231–236.
- Zanelli, C. I., Hennige, C. W., and Sanghvi, N. T. (1994). "Design and characterization of a 10 cm annular array transducer for high intensity focused ultrasound (HIFU) applications," *Proc.-IEEE Ultrason. Symp.* **3**, 1887–1890.

PROGRAM OF

The 149th Meeting of the Acoustical Society of America

Vancouver, Canada • Hyatt Regency Vancouver • 16–20 May 2005

1a MON. AM

NOTE: All Journal articles and Letters to the Editor are peer reviewed before publication. Program abstracts, however, are not reviewed before publication, since we are prohibited by time and schedule.

MONDAY MORNING, 16 MAY 2005

REGENCY E, 7:40 A.M. TO 12:00 NOON

Session 1aAA

Architectural Acoustics, Education in Acoustics, Noise and Psychological and Physiological Acoustics: Topical Meeting on Classroom Acoustics—The Research Perspective I

Murray R. Hodgson, Cochair

*Univ. of British Columbia, School of Occupational and Environmental Hygiene, 2206 East Mall,
Vancouver, BC V6T 1Z3, Canada*

Lily M. Wang, Cochair

Univ. of Nebraska Lincoln, Architectural Engineering, Peter–Kiewit Institute, 1110 South 67th St., Omaha, NE 68182-0681

Chair's Introduction—7:40

Invited Papers

7:45

1aAA1. Implications of the road traffic and aircraft noise exposure and children's cognition and health (RANCH) study results for classroom acoustics. Stephen A Stansfeld, Charlotte Clark, and on behalf of the RANCH Study Team (Ctr. for Psychiatry, Barts and the London, Queen Marys School of Medicine and Dentistry, Mile End Rd., London, E1 4NS, UK)

Studies in West London have found associations between aircraft noise exposure and childrens' cognitive performance. This has culminated in the RANCH Study examining exposure-effect associations between aircraft and road traffic noise exposure and cognitive performance and health. The RANCH project, the largest cross-sectional study of noise and childrens health, examined 2844 children, 9–10 years old, from 89 schools around three major airports: in the Netherlands, Spain and the United Kingdom. Children were selected by external aircraft and road traffic noise exposure at school predicted from noise contour maps, modeling and on-site measurements. A substudy indicated high internal levels of noise within classrooms. Schools were matched for socioeconomic position within countries. Cognitive and health outcomes were measured by standardized tests and questionnaires administered in the classroom. A parental questionnaire collected information on socioeconomic position, parental education and ethnicity. Linear exposure-effect associations were found between chronic aircraft noise exposure and impairment of reading comprehension and recognition memory, maintained after adjustment for mothers education, socioeconomic factors, longstanding illness and classroom insulation. Road traffic noise exposure was linearly associated with episodic memory. The implications of these results for childrens' learning environments will be discussed. [Work supported by European Community (QLRT-2000-00197) Vth framework program.]

8:15

1aAA2. The probability of young children understanding their teacher in everyday teaching situations. John S. Bradley and Hiroshi Sato (Inst. for Res. in Construction, Natl. Res. Council, Montreal Rd., Ottawa, Canada K1A 0R6)

This paper examines the probability of grades 1, 3, and 6 students understanding their teacher in a normal teaching activity. The results of speech intelligibility tests in 34 enclosed classrooms were analyzed to relate childrens' ability to understand speech to measured signal-to-noise ratios (as well as Speech Transmission Index and useful-to-detrimental sound ratios) for each age group. Measurements of speech and noise levels during a normal teaching activity were used to determine the distribution of signal-to-noise

ratios in normal teaching situations. Combining these two sets of data shows that while 51% of the grade 6 students experienced near-to-ideal acoustical conditions only 9% of the grade 1 students had close-to-ideal conditions in these apparently good classrooms. Estimates of ideal goals for classroom signal-to-noise ratios for each age group will be presented. [Work supported by CLLRnet.]

8:35

1aAA3. Speech intelligibility tests with real and simulated classroom noise sources. Bumjun Kim and Gary Siebein (Univ. of Florida School of Architecture, P.O. Box 115702, Gainesville, FL 32607)

Speech intelligibility tests using the MRT stimuli were administered to 13 college age students in actual classroom spaces in the presence of noise sources identified in a survey of 41 elementary school classrooms in 7 schools. Twenty two different noise sources were identified during the survey of schools in a local school district. A-weighted, C-weighted, octave band and 1/3 octave band measurements of 182 individual noise sources were recorded in the sampled rooms. Noise sources were categorized as students speaking, HVAC equipment, other building equipment, water running in sinks, computers, lighting fixtures, A/V equipment, and miscellaneous sources. The noise source measurements provide a useful data base for future classroom acoustical studies. Five noise levels (35, 45, 55, 65, and 75 dBA) of 4 different types of noise were tested in the rooms. The speech intelligibility tests found significant effects (at the 0.01 level) from level, spectra, type, content, and annoyance caused by the different noise sources. A predictive model based on these variables had an R2 of 0.90 for speech intelligibility under the extremely diverse set of conditions tested.

8:55

1aAA4. Auralization study of optimum reverberation for speech intelligibility for normal and hearing-impaired listeners. Wonyoung Yang, Murray Hodgson (School of Occupational and Environ. Hygiene, Univ. of British Columbia, 3rd Fl., 2206 East Mall, Vancouver, BC, Canada V6T1Z3), and Maki Ezaki (Central West Health Co. Speech-Lang. Pathol. and Audiol., Grand Falls-Windsor, NF, Canada A2A 2E1)

Reverberation and signal-to-noise level difference are two major factors affecting speech intelligibility. They interact in rooms. Past work has accounted for noise using a constant received background-noise level. Noise is actually generated by sources, and varies, and affects speech intelligibility differently, throughout the classroom, depending on where the sources are located. Here, a speech-babble noise source located at different positions in the room was considered. The relative output levels of the speech and noise sources, resulting in different signal-to-noise level differences, were controlled, along with the reverberation. The binaural impulse response of a virtual idealized classroom model was convolved with the Modified Rhyme Test (MRT) source and babble-noise signals in order to find the optimal configuration for speech intelligibility. Speech-intelligibility tests were performed with normal and hard-of-hearing subjects in each of 16 conditions which were combinations of reverberation time, signal-to-noise level difference, and speech- and noise-source locations. For both normal and hearing-impaired subjects, when the speech source was closer to the listener than the noise source, the optimal RT was zero. When the noise source was closer to the listener than the speech source, the optimal RT was generally non-zero. This agrees with theoretical results.

9:15

1aAA5. Optimal speech level for speech transmission in a noisy environment for young adults and aged persons. Hayato Sato, Ryo Ota, Masayuki Morimoto (Environ. Acoust. Lab., Kobe Univ., Nada, Kobe 657-8501, Japan, hayato@kobe-u.ac.jp), and Hiroshi Sato (Natl. Inst. of Adv. Industrial Sci. and Technol., Tsukuba, Ibaraki 305-8566, Japan)

Assessing sound environment of classrooms for the aged is a very important issue, because classrooms can be used by the aged for their lifelong learning, especially in the aged society. Hence hearing loss due to aging is a considerable factor for classrooms. In this study, the optimal speech level in noisy fields for both young adults and aged persons was investigated. Listening difficulty ratings and word intelligibility scores for familiar words were used to evaluate speech transmission performance. The results of the tests demonstrated that the optimal speech level for moderate background noise (i.e., less than around 60 dBA) was fairly constant. Meanwhile, the optimal speech level depended on the speech-to-noise ratio when the background noise level exceeded around 60 dBA. The minimum required speech level to minimize difficulty ratings for the aged was higher than that for the young. However, the minimum difficulty ratings for both the young and the aged were given in the range of speech level of 70 to 80 dBA of speech level.

9:35

1aAA6. Acoustics and sociolinguistics: Patterns of communication in hearing impairing classrooms. William McKellin, Kimary Shahin (Dept. of Anthropology and Sociology, Univ. of British Columbia, Vancouver, BC, Canada V6S 1Z1), Janet Jamieson (Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z1), Murray Hodgson (UBC Acoust. and Noise Res. Group, SOEH, Vancouver, BC, Canada V6T 1Z3), and Kathleen Pichora-Fuller (Univ. of Toronto at Mississauga, Mississauga, ON, Canada L5L 1C6)

In elementary school classes, noise during student led activities is often taken as evidence of successful interaction and learning. In this complex social environment of elementary school classrooms, acquisition of complex language and social skills—the focus of activities in early education—is expected to take place in hearing-hostile environments. Communication and language processing in these contexts requires interactive strategies, discourse forms, and syntactic structures different from the educationally desired forms used in acoustically advantageous environments. Recordings were made of the interaction of groups of students in grades 1–3, 5, and 7 during collaborative group work in their regular classrooms. Each student wore microphones at the ear level and head-mounted video cameras. Each group as a whole was also audio- and videotaped and noise level readings were recorded. Analysis of the acoustical and phonological properties of language heard by each student has demonstrated that the language variety used in these noisy and reverberant settings is similar to that of individuals with hearing impairments. This paper reports similarities between the syntactic structures and pragmatic strategies used by hearing impaired children and normally hearing children in noisy contexts. [Work supported by Peter Wall Institute for Advanced Studies, University of British Columbia.]

1aAA7. Social and academic implications of acoustically hostile classrooms for hard of hearing children. Janet R. Jamieson (Faculty of Education/ECPS, Univ. of British Columbia, 2125 Main Mall, Vancouver, BC, Canada V6T 1Z4)

The correlation between lowered academic achievement and classroom noise has been demonstrated for normally hearing children (Shield and Dockrell, 2003). However, the implications of poor classroom acoustics on the socialization and academic performance of children who are hard of hearing have not been examined. Eleven hard of hearing students in one school district, ranging from kindergarten to grade 7, were the foci of the present study. Acoustic measurements of each of the 11 classrooms in both unoccupied and occupied conditions revealed that all classrooms were acoustically challenging for the hard of hearing students, particularly at transition times, when ventilation was operational, and in the primary grades, when language learning needs are greatest. Interviews with parents and teachers underscored the difficulty these students experienced in comprehending teacher instructions and participating in group work. The students seldom initiated conversation or seatwork independently, but, rather, followed the lead of their peers. The hard of hearing students experienced difficulties in understanding or participating in informal peer-to-peer conversations in the classroom, and parents and teachers attributed the children's frequent social isolation and withdrawal at school to the combined effects of poor hearing abilities and hostile classroom acoustics. [Work supported by Hampton Research Fund.]

10:15–10:30 Break

10:30

1aAA8. Classroom acoustics: Three pilot studies. Joseph J. Smaldino (Northern Illinois Univ., De Kalb, IL 60115)

This paper summarizes three related pilot projects designed to focus on the possible effects of classroom acoustics on fine auditory discrimination as it relates to language acquisition, especially English as a second language. The first study investigated the influence of improving the signal-to-noise ratio on the differentiation of English phonemes. The results showed better differentiation with better signal-to-noise ratio. The second studied speech perception in noise by young adults for whom English was a second language. The outcome indicated that the second language learners required a better signal-to-noise ratio to perform equally to the native language participants. The last study surveyed the acoustic conditions of preschool and day care classrooms, wherein first and second language learning occurs. The survey suggested an unfavorable acoustic environment for language learning.

10:50

1aAA9. Classroom noise and children learning in a second language. Peggy Nelson, Kathryn Kohnert, Sabina Sabur (Dept. of Speech-Lang.-Hearing Sci., Univ. of Minnesota, 164 Pillsbury Dr. SE, Minneapolis, MN 55455), and Daniel Shaw (Jefferson Community School, Minneapolis, MN 55405)

The presence of background noise affects children more negatively than adults. Understanding speech in noise is a skill that continues to develop well into a child's adolescent years. Children's experience with a specific language also may affect their ability to make sense of incoming speech. Research suggests that even for adults the presence of background noise negatively affects the ability to listen in a second language. Two studies were conducted to investigate the effects of classroom noise on attention and speech perception in native Spanish-speaking second graders learning English as their second language (L2), as compared to English-only speaking peers (EO). In Study 1 we measured childrens' on-task behavior during instructional activities with and without soundfield amplification. In Study 2 we measured the effects of noise (+10 dB signal-to-noise ratio) using an experimental English word-recognition task. Findings indicate although there were no effects of amplification on on-task behavior, word-recognition performance declined significantly for both EO and L2 groups in the noise condition. In particular, the impact of the noise was disproportionately greater for the L2 group. Children learning in their L2 appear to be at a distinct disadvantage when listening in rooms with typical noise and reverberation.

11:10

1aAA10. Effects of road traffic and aircraft noise upon children's academic attainments. Bridget Shield (Dept. of Eng. Systems, Faculty of Eng., Sci. and Built Environment, London South Bank Univ., London SE1 0AA, UK), Julie Dockrell (London Univ., London WC1H 0AA, UK), and Gael Vilatarsana (London South Bank Univ., London SE1 0AA, UK)

The effects of environmental noise upon the academic performance of children aged 7 and 11 years in primary schools in London (UK) have been investigated. Noise surveys were carried out to measure levels of environmental noise during the school day outside 175 schools across London. The majority of the schools were in densely populated areas within 5 miles of central London, where road traffic was the dominant noise source. Thirty three of the schools were in a less densely populated area to the west of London near Heathrow Airport, and were subject to predominantly aircraft noise. The noise levels measured outside each school have been correlated with the results of standard tests in Reading, Writing, Mathematics, English, and Science, which are taken by all children aged 7 and 11 in England and Wales. Significant negative correlations were found between noise levels and many of the test scores, the correlations being stronger in the central London areas than in the schools around Heathrow. These results show that environmental noise has a detrimental effect upon childrens' academic performance, the effect remaining apparent when data were corrected for socio-economic factors such as social deprivation.

11:30

1aAA11. Acoustic environment challenges for the unique communication conditions in group learning classes in elementary school classrooms. Louis Sutherland (27803 Longhill Dr., Rancho Palos Verdes, CA 90275), David Lubman (Westminster, CA 92683), and Karl Pearsons (Woodland Hills, CA 91364)

Unlike the traditional “sage-on-the-stage” configuration of many K–12 classrooms, the group learning or “guide-on-the-side” configuration does not involve communication between a teacher in front of a seated class of 20 to 30 students. Instead, it can involve, most of the time, communication between the teacher and each of several small groups of students interacting, aurally, with each other. To maintain the desired 15 dB signal-to-noise ratio intended as the rationale for the ANSI standard, S12.60-2002 on classroom acoustics, the “noise” heard by participants in one of the groups is likely to include the speech levels generated by the participants in the other groups as well as the background noise in the unoccupied classroom. Thus, specification of the speech level within (i.e. the “signal”), and between (i.e. part of the “noise”) the learning groups, must be considered. Data available to evaluate these speech levels are reviewed and possible models considered to account for the Lombard effect for voice levels of both the teacher and the students. Some of the gaps in these data are suggested as a challenge to stimulate further studies on speech levels of teachers and students in a wide range of communication conditions.

11:45

1aAA12. Hearing impaired speech in noisy classrooms. Kimiry Shahin (Dept. of English Lang. & Translation, Effat College of Al-Faisal Univ., Jeddah 21478, Saudia Arabia), William H. McKellin, Janet Jamieson, Murray Hodgson (Univ. of British Columbia, Vancouver, Canada V6T 1Z3), and M. Kathleen Pichora-Fuller (Univ. of Toronto, Mississauga, Canada L5L 1C6)

Noisy classrooms have been shown to induce among students patterns of interaction similar to those used by hearing impaired people [W. H. McKellin *et al.*, GURT (2003)]. In this research, the speech of children in a noisy classroom setting was investigated to determine if noisy classrooms have an effect on students’ speech. Audio recordings were made of the speech of students during group work in their regular classrooms (grades 1–7), and of the speech of the same students in a sound booth. Noise level readings in the classrooms were also recorded. Each student’s noisy and quiet environment speech samples were acoustically analyzed for prosodic and segmental properties (f_0 , pitch range, pitch variation, phoneme duration, vowel formants), and compared. The analysis showed that the students’ speech in the noisy classrooms had characteristics of the speech of hearing-impaired persons [e.g., R. O’Halpin, *Clin. Ling. and Phon.* **15**, 529–550 (2001)]. Some educational implications of our findings were identified. [Work supported by the Peter Wall Institute for Advanced Studies, University of British Columbia.]

MONDAY MORNING, 16 MAY 2005

PLAZA A, 8:00 TO 11:35 A.M.

Session 1aAO

Acoustical Oceanography and Underwater Acoustics: Riverine Acoustics I

Ken Cooke, Chair

Biological Sciences Branch, Fisheries and Oceans, Pacific Biological Station, Nanaimo, BC V9R 5K6, Canada

Chair’s Introduction—8:00

Invited Papers

8:05

1aAO1. Fisheries management applications of riverine hydroacoustics: 30 years’ experience with applied technology in the practical arena. Michael F. Lapointe (Pacific Salmon Commission, 600-1155 Robson St., Vancouver, BC, Canada V6E 1B5, lapointepsc.org)

Some examples of the successes and challenges encountered by the Pacific Salmon Commission in the application of riverine hydroacoustics to fisheries management of Fraser River sockeye salmon are reviewed. Riverine hydroacoustics estimates have been an integral part of the fisheries data collected by the Pacific Salmon Commission for over 30 years. Real time estimates of fish passage provide intra-seasonal feedback on the progress toward escapement targets and information about changing total abundance levels. This information has allowed managers to adjust fisheries schedules and improved their ability to meet catch and escapement objectives. Despite these successes, application of technology has encountered a number of challenges including: (1) the interpretation of acoustics data in determining fish targets, (2) quantification of accuracy of hydroacoustic estimates in large rivers, (3) misperceptions about estimation methods by the public, and (4) inevitable comparisons with estimates from other sources and their effect on perceived accuracy of the hydroacoustic estimates. Lessons learned from the Pacific Salmon Commission experience are summarized with the objective of helping others engaged in the application of riverine acoustics technology to fisheries management problems.

8:35

1aAO2. A retrospective on hydroacoustic assessment of fish passage in Alaskan rivers. Debby Burwen, Steve Fleischman, Suzanne Maxwell, and Carl Pfisterer (Alaska Dept. of Fish and Game, 333 Raspberry Rd., Anchorage, AK 99518)

The Alaska Department of Fish and Game (ADFG) has enumerated fish stocks in rivers for over 30 years using a variety of acoustic technologies including single-, dual-, and split-beam sonar. Most recently, ADFG has evaluated a relatively new sonar technology at several sites in Alaska to determine its applicability to counting migrating fish in rivers. The new system, called a Dual

frequency IDentification SONar (DIDSON), is a high-definition imaging sonar designed and manufactured by the University of Washington's Applied Physics Lab for military applications such as diver detection and underwater mine identification. Results from experiments conducted in 2002–2004 indicate that DIDSON provides significant improvements in our ability to detect, track, and determine the direction of travel of migrating fish in rivers. One of the most powerful uses of the DIDSON has been to combine its camera-like images of fish swimming behavior with corresponding split-beam data. These linked datasets have allowed us to evaluate the effects of fish orientation and swimming behavior on echo shape parameters that have proven useful in the classification of certain fish species.

9:05

1aAO3. Split-beam sonar observations of targets as an aid in the interpretation of anomalies encountered while monitoring migrating adult salmon in rivers. George M. W. Cronkite (Dept. of Fisheries and Oceans, Pacific Biological Station, 3190 Hammond Bay Rd., Nanaimo, BC, V9T 6N7, Canada, cronkiteg@pac.dfo-mpo.gc.ca), Hermann J. Enzenhofer (Cultus Lake Res. Lab., Cultus Lake, BC, V2R 5B6, Canada), and Andrew P. Gray (Pacific Salmon Commission, Vancouver, BC, V6E 1B5, Canada)

The experiments described in this paper relate known target configurations under controlled conditions to acoustic characteristics of multiple moving fish. This was done to increase the understanding of the interactions between targets and the effects these interactions have on the measurement of the number of salmon migrating in rivers. Multiple targets in various configurations were passed through a horizontally oriented 4×10 beam from a split-beam echo sounder. The effects on measurements of target strength, detection probability and target location in the beam are presented. There was a reduction in target detection due to the single-target selection criteria implemented by the hydroacoustic system. The conditions in a river were mimicked to demonstrate how a close range fish target may modify the beam geometry allowing detection of previously undetected targets. The effects of moving targets into radial alignment were demonstrated along with shadowing conditions that can cause extinction of target echoes.

9:25

1aAO4. Acoustic measurement of the behavioral response of Arctic riverine fish to seismic sound. Eric Gyselman and John Jorgenson (Fisheries and Oceans Canada, 501 Univ. Crescent, Winnipeg, MB, Canada, gyselmane@dfo-mpo.gc.ca)

Renewed interest in oil and gas development in the Canadian Arctic has led to proposals to conduct seismic surveys along the entire length of the Mackenzie River. However, little is known about the effects of seismic sound (air guns) on fish in riverine environments. In 2004, Fisheries and Oceans Canada carried out a study to look at the effects of seismic sound on the physiology and behavior of fish in the Mackenzie River. The behavioral component used a split-beam acoustic system to measure the response of fish to varying levels of seismic sound. Targets were tracked with SonarDats Echoview Tracking Module. Two experiments were carried out. In the first, the acoustic launch was anchored over a concentration of fish while the seismic barge approached. This experiment simulated the conditions proposed for the actual seismic survey. During the second, the acoustic launch was allowed to drift over concentrations of fish. The seismic barge was stationary. When individual fish were seen in the acoustic beam, the air guns were fired. This experiment measured the overt flight response of fish to the seismic sound. Initial results from both studies indicate that fish show no direct behavioral response to the seismic sound.

Contributed Paper

9:45

1aAO5. Innovative techniques for analyzing the three-dimensional behavioral results from acoustically tagged fish. Tracey W. Steig and Mark A. Timko (Hydroacoustic Technol., Inc., 715 N.E. Northlake Way, Seattle, WA 98105)

Acoustic tags were used to monitor the swimming patterns of downstream migrating salmon smolts approaching various dams on the Columbia River, USA. Downstream migrating yearling chinook (*Oncorhynchus tshawytscha*), steelhead (*Oncorhynchus mykiss*), sockeye (*Oncorhynchus nerka*), and sub-yearling chinook smolts were surgically implanted with acoustic tags. Fish were tracked in three-dimensions as they approached

and passed into the turbine intakes, spillways, and surface bypass channel entrances at the dams during the 2004 spring and summer outmigrations. A number of advances in the analysis techniques and software have been made over the past few years. Some of these improvements include the development of various fish density algorithms, stream trace modeling analysis, and advances of three-dimensional animation programs. Three-dimensional tracks of fish approaching the turbine intakes, spillways, and surface bypass channel entrances will be presented. Concentrations of fish passage will be presented as three-dimensional fish densities superimposed over dam structures. Stream trace modeling animation will be presented showing predicted fish passage routes.

10:00–10:15 Break

Invited Paper

10:15

1aAO6. Salmon enumeration in the Fraser River with the dual-frequency identification sonar (DIDSON) acoustic imaging system. John A. Holmes, George Cronkite (Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Rd., Nanaimo, BC, V9T 6N7, Canada), and Hermann J. Enzenhofer (Fisheries and Oceans Canada, Cultus Lake, BC V2R 5B6, Canada)

Reliable data on the number of salmon entering tributaries of the Fraser River to spawn (escapement) is needed for Pacific salmon management. Existing escapement techniques are costly and the number of populations requiring assessments has risen because of stock rebuilding efforts. The efficacy of a DIDSON acoustic imaging system for salmon stock assessment was investigated. Sixteen potential sites within the Fraser watershed were surveyed and based on channel morphology, bottom morphology, flow pattern, fish behavior and location relative to spawning grounds, ten sites in six rivers meet the needs of fisheries managers and the DIDSON

system for escapement estimates. Fish count data from the DIDSON were compared to data from a counting fence (used as a standard) using regression techniques, resulting in relationships with slopes ranging from 0.98 to 1.02. The precision of DIDSON counts >50 (measured by CV) among three readers was 1.7%. This work supports the conclusion that the DIDSON system can deliver escapement estimates whose accuracy, precision and scientific defensibility is consistent with or better than existing escapement techniques and at a lower operating cost to assessment programs. [Work supported by the Southern Boundary Restoration and Enhancement Fund of the Pacific Salmon Commission.]

Contributed Papers

10:35

1aAO7. Survey of dual frequency identification sonar (DIDSON) applications in fisheries assessment and behavioral studies. Edward O. Belcher (Sound Metrics Corp., 6824 NE 160th St., Kenmore, WA 98028)

The Dual Frequency Identification Sonar (DIDSON) is a forward-looking sonar that operates in shallow riverine environments with rocky, uneven substrates and near concrete structures such as dams. This allows a number of fisheries applications in environments previously too hostile for reliable sonar operation. Currently 35 DIDSONs have been obtained by 16 groups to accomplish a variety of fish assessment and behavioral studies. This paper surveys the work of these groups and highlights novel assessments allowed by this new acoustic tool. The groups include Alaska Department of Fish and Game, U.S. Fish and Wildlife Service, NOAA, USGS, Bureau of Reclamation, California Department of Water Resources, Pacific Northwest National Laboratories, Nez Perce Tribal Fisheries, Puyallup Tribal Fisheries, Department of Fisheries and Oceans Canada, and Fisheries Engineering Japan. Assessments include: (1) Counting fish migrating up rivers of various sizes, bottom substrates, turbidity, and velocity; (2) Analysis of fish behavior at (A) prototype fish protection devices on dams, (B) irrigation intakes along muddy rivers, (C) intakes of trawl nets; and (3) Detection and measurement of redds in alluvial riverbeds. [Work for the survey supported by Sound Metrics Corp.]

10:50

1aAO8. Differentiating fish targets from non-fish targets using an imaging sonar and a conventional sonar: Dual frequency identification sonar (DIDSON) versus split-beam sonar. Yunbo Xie, Andrew P. Gray, and Fiona J. Martens (Pacific Salmon Commission, 600-1155 Robson St., Vancouver, BC, Canada V6E 1B5, xie@psc.org)

A key requirement in applying acoustic techniques to estimating fish abundance is the removal of non-fish targets from the database. In a riverine environment, debris, entrained air bubbles, bottom objects are common ambient targets which can effectively scatter the probing sound from a fisheries sonar system, and cause strong echoes for the system. A conventional sonar system provides limited and highly simplified information for a detected target, which results in difficulty in separating fish from other targets. Recently developed DIDSON sonar utilizes imaging sonar technology to provide photo-quality images of underwater objects, making possible the visual interpretation of targets. A DIDSON system was deployed in the Fraser River at Mission, British Columbia during the salmon migration in 2004. Data were collected simultaneously from the DIDSON sonar and from a 200-kHz split-beam sonar. These data allow for comparisons of estimates of upstream salmon flux acquired concurrently by the imaging and the split-beam sonar systems.

11:05

1aAO9. Echoview software for tracking and counting fish detected with a dual frequency identification sonar (DIDSON) imaging sonar: A flexible approach to signal enhancement and object detection prior to target tracking. Ian Higginbottom (SonarData Pty Ltd., G.P.O. Box 1387, Hobart, Tasmania 7001, Australia)

The data from the DIDSON (Dual frequency IDentification SONar) are of startling resolution, and show swimming fish in unprecedented detail. Although the fish objects are visually striking, object detection must be achieved in post processing and is a challenge when compared with conventional split beam systems. Data processing can be applied to automatically remove static objects and noise and to enhance the signal from moving objects in DIDSON data. Moving objects are then detected as targets within individual DIDSON frames and converted to $X-Y-Z$ -time position data suitable for input to a tracking algorithm. The initial data manipulation for noise removal and object detection is achieved through the application of successive modular operators. The results of each operation can be viewed as a "virtual variable" enabling each step to be scrutinized and optimized by the operator. This modular approach to data manipulation, target detection, and fish tracking will enable the rapid development of new techniques and tools. It is a "future-proof" data processing solution.

11:20

1aAO10. Integrating new technologies into split-beam riverine sonar systems. James Dawson (BioSonics, Inc., 4027 Leary Way NW, Seattle, WA 98107)

While development has continued on both acoustic hardware and data processing software, parallel development has involved integration of other technologies. Users may now access and control acoustic systems over the Internet or phone lines. Smart system software can perform self-diagnostics, alert remote users of system status, and automatically adjust operational and physical parameters (such as transducer orientation) during autonomous operation. In addition, advanced remote and autonomous systems are capable of real-time analysis and decision-making functions enabling near real-time decisions and automated interactions with a variety of devices, such as modems, Ethernet adapters, motor controls, and on-site electro-mechanical devices. Details and benefits of these advances are presented in a variety of case studies.

Session 1aBB

Biomedical Ultrasound/Bioresponse to Vibration and Physical Acoustics: Tissue Response to Shock Waves I

Michael R. Bailey, Cochair

Applied Physics Lab., Center for Industrial and Medical Ultrasound, 1013 NE 40th St., Seattle, WA 98105

Reiner Schultheiss, Cochair

*SWS Shock Wave Systems AG, Wilen 4, 8574 Illighausen, Switzerland***Chair's Introduction—8:00****Invited Papers****8:05****1aBB1. Extracorporeal shock wave therapy for non-unions and delayed fracture healing.** Wolfgang Schaden, Andreas Fischer, Andreas Sailler, and Ender Karadas (Landstrasser Hauptstrasse 83, Vienna 1030, Austria, ismst@aon.at)

Although the primary management of fractures is highly developed in Central Europe 1% of fractures develop a non-union. After successful pilot studies the Traumacenter Meidling started in December 1998 to treat non-unions regularly with shock wave therapy. From December 1998 to August 2004, 1153 patients with non-union and delayed healing fractures were treated. The results of 755 patients are available up to September 2004. The patients consisted of 250 (33%) female and 505 (67%) male. The mean age was 44.1 years (10; 90). The mean age of the non-union was 15.5 months. In 74 (10%) osteomyelitis was present before shockwave therapy. Out of 755 non-unions 593 (79%) achieved bony healing. As expected, the subgroup of 284 delayed unions (shockwave therapy 3–6 months after the trauma or the last surgery concerning the bone) showed the best results. 245 (86%) healed. Out of 471 non-unions being older than 6 months 348 (72%) achieved bony healing. Because of the efficacy and the lack of complications as well as the economic advantage in comparison to surgery, shockwave therapy is considered as therapy of first choice in the treatment of non-union and delayed healing fractures.

8:20**1aBB2. The effect of shock wave treatment at the tendon-bone interface.** Ching-Jen Wang (Chang Gung Memorial Hospital at Kaohsiung, Taiwan, 123, Ta Pei Rd., Niao Sung Hsiang, Kaohsiung 833, Taiwan)

This study was performed to investigate the effect of shock wave treatment on the healing at tendon-bone interface. Thirty-six New Zealand White rabbits were used in this study. The anterior cruciate ligament was excised and replaced with the long digital extensor. The right knees (study group) were treated with 500 impulses of shock waves at 14 kV, while the left knees (control group) received no shock waves. Twenty-four rabbits were sacrificed at 1, 2, 4, 8, 12 and 24-week intervals. The specimens were studied with histomorphological examination and immunohistochemical stains for neovascularization and angiogenic growth factors. Twelve rabbits were sacrificed at 12 and 24 weeks for biomechanical analysis. The results demonstrated that the study group showed significantly more trabecular bone around the tendon and better bonding between bone and tendon as compared with the control group. The expressions of angiogenic growth factors were significantly higher in the study group than the control group. The tensile strength of the tendon-bone interface was significantly higher in the study group than the control group. In conclusion, shock wave treatment significantly improves the healing of the tendon-bone interface in a bone tunnel in rabbits. The effect of shock waves appears to be time-dependent.

8:35**1aBB3. Extracorporeal shock wave therapy in orthopedics, basic research, and clinical implications.** Joerg Hausdorf, Volkmar Jansson, Markus Maier (Orthopedic Dept., Ludwig-Maximilians-Univ. Munich, Marchioninstr. 15, 81377 Munich, Germany, joerg.hausdorf@med.uni-muenchen.de), and Michael Delius (Ludwig-Maximilians-Univ. Munich, 81377 Munich, Germany)

The molecular events following shock wave treatment of bone are widely unknown. Nevertheless patients with osteonecrosis and non unions are already treated partly successful with shock waves. Concerning the first indication, the question of the permeation of the shock wave into the bone was addressed. Therefore shockwaves were applied to porcine femoral heads and the intraosseous pressure was measured. A linear correlation of the pressure to the intraosseous distance was found. Approximately 50% of the pressure are still measurable 10 mm inside the femoral head. These findings should encourage continued shock wave research on this indication. Concerning the second indication (non union), osteoblasts were subjected to 250 or 500 shock waves at 25 kV. After 24, 48, and 72 h the levels of the bone and vascular growth factors bFGF, TGFbeta1, and VEGF were examined. After 24 h there was a significant increase in bFGF levels ($p < 0.05$) with significant correlation ($p < 0.05$) to the number of impulses. TGFbeta1, and VEGF showed no significant changes. This may be one piece in the cascade of new bone formation following shock wave treatment and may lead to a more specific application of shock waves in orthopedic surgery.

1aBB4. Experience with extracorporeal shock wave therapy (ESWT) in the United States. John P. Furia (SUN Orthopedics and Sports Medicine, 900 Buffalo Rd., Lewisburg, PA 17837)

The purpose of this presentation is to summarize the literature and to report on single treatment, high-energy ESWT for the treatment of chronic plantar fasciitis and lateral epicondylitis. Fifty-three patients (60 heels) were treated with 3800 shock waves. Sixteen patients (19 heels) were active, 21 (22 heels), were moderately active, and 16 (19 heels) were sedentary. Twelve weeks post treatment, mean visual analog scores (VAS) for the entire group improved from 9.2 to 2.4 ($p < 0.05$), RAND-Physical Functioning score improved from 40.4 to 91.5 ($p < 0.05$), and RAND-Pain score improved from 33.3 to 90 ($p < 0.05$). Fifty heels (83.3%) were assigned an excellent or good result. Thirty-six patients with chronic lateral epicondylitis were treated with 3200 shock waves. There were 9 workers compensation and 27 non-workers compensation patients. Twelve weeks post treatment, the mean VAS for the entire group improved from 8.0 to 2.5 ($p < 0.05$), and the mean RAND-Physical Functioning score improved from 65.6 to 88.0 ($p < 0.05$). Twenty-eight elbows (77.8%) were assigned an excellent or good result. In both trials, outcome was similar for each subgroup. There were no significant complications in either trial. Using the therapeutic parameters applied, ESWT is a safe and effective treatment for chronic plantar fasciitis and lateral epicondylitis.

1aBB5. Adverse effects of shock waves and strategies for improved treatment in shock wave lithotripsy. James A. McAteer, Andrew P. Evan, Bret A. Connors, James C. Williams, Jr. (Dept. of Anatomy and Cell Biol., Indiana Univ. School of Medicine, 635 Barnhill Dr., Indianapolis, IN 46202-5120, mcaateer@anatomy.iupui.edu), and Lynn R. Willis (Indiana Univ. School of Medicine, Indianapolis, IN 46202-5120)

Lithotripter SWs rupture blood vessels in the kidney. This acute trauma, accompanied by a fall in renal function, can lead to significant long-term effects such as profound scarring of the kidney cortex and renal papillae permanent loss of functional renal mass. SWL has been linked to new-onset hypertension in some patients, and recent studies suggest that multiple lithotripsies can actually alter a patient's stone disease leading to formation of stones (brushite) that are harder to break. Cavitation and shear appear to play a role in stone breakage and tissue damage. Progress in understanding these mechanisms, and the renal response to SWs, has led to practical strategies to improve treatment. Slowing the SW-rate, or initiating treatment at low kV/power both improve stone breakage and reduce the number of potentially tissue-damaging SWs needed to achieve comminution. The observation that SWs cause transient vasoconstriction in the kidney has led to studies in pigs showing that a pre-conditioning dose of low-energy SWs significantly reduces trauma from subsequent high-energy SWs. Thus, SWs can induce adverse effects in the kidney, but what we have learned about the mechanisms of SW action suggests strategies that could make lithotripsy safer and more effective. [Work supported by NIH-DK43881, DK55674.]

1aBB6. In vitro comparison of shock wave lithotripsy machines. Joel M. Teichman (St. Paul's Hospital, Burrard Bldg. C307, 1081 Burrard St., Vancouver, BC, Canada V6Z 1Y6, jteichman@providencehealth.bc.ca), Patricia P. Ceconi (Univ. of Texas Health Sci. Ctr., San Antonio, TX 78229), Margaret S. Pearle (Univ. of Texas Southwestern Medical Ctr., Dallas, TX), and Ralph V. Clayman (Univ. of California, Irvine, CA)

We tested the hypothesis that shock wave lithotripsy machines vary in the ability to fragment stones to small size. Calcium oxalate monohydrate, calcium phosphate, cystine and struvite calculi were fragmented *in vitro* with the Dornier HM3, Storz Modulith SLX, Siemens Lithostar C, Medstone STS-T, HealthTronics LithoTron 160, Dornier Doli S and Medispec Econolith lithotriptors. Stones were given 2000 shocks or the FDA limit. Post-lithotripsy fragment size was compared. Struvite calculi were completely fragmented by all devices. The mean incidence of calcium phosphate dihydrate, calcium oxalate monohydrate, and cystine stones rendered into fragments greater than 2 mm was 0% for the HM3, Modulith SLX and Lithostar C, 10% for the STS-T, 3% for the LithoTron 160, 29% for the Doli and 18% for the Econolith ($p = 0.04$); 0% for the HM3, Modulith SLX, Lithostar C, STS-T and LithoTron 160, 4% for the Doli and 9% for the Econolith ($p = 0.15$); 1% for the HM3, 0% for the Modulith SLX, 1% for the Lithostar C, 10% for the STS-T, 14% for the LithoTron 160, 3% for the Doli and 9% for the Econolith ($p = 0.44$), respectively. Shock wave lithotriptors vary in fragmentation ability.

1aBB7. Shock wave lithotripsy at 60 or 120 shocks per minute: A randomized, double-blinded trial. Kenneth Pace, Daniela Ghiculete, Melanie Harju, and R. John Honey (St. Michael's Hospital, Univ. of Toronto, 61 Queen St. E, Ste. 9-106, Toronto, ON, Canada M5C 2T2)

Rate of shock wave administration is a factor in the per-shock efficiency of SWL. Decreasing shock wave frequency from 120 shocks per minute (s/m) may improve stone fragmentation. This study is the first to test this hypothesis *in vivo*. Patients with previously untreated radio-opaque kidney stones were randomized to SWL at 60 or 120 s/m and followed at 2 weeks and 3 months. Primary outcome was success rate, defined as stone-free or asymptomatic fragments 5 mm in size 3 months post-treatment. 111 patients were randomized to 60 s/m and 109 to 120 s/m. The groups were comparable on age, gender, BMI, stent status, and initial stone area. Success rate was higher for 60 s/m (75% versus 61%, $p = 0.027$). Patients with stone area 100 mm² experienced the greatest benefit: success rates were 71% for 60 s/m versus 32% ($p = 0.002$), and stone-free rates were 60% versus 28% ($p = 0.015$). Repeat SWL treatment was required in 32% treated at 120 s/m versus 18% ($p = 0.018$). Fewer shocks were required (2423 versus 2906, $p = 0.001$), but treatment time was longer (40.6 versus 24.2 minutes, $p = 0.001$). SWL treatment at 60 s/m yields better outcomes than 120 s/m, particularly for stones 100 mm².

1aBB8. Cavitation, membrane transfer, and molecular response. Michael Delius and Gerhard Adams (Inst. for Surgical Res., Klinikum Grosshadern, 81366 Munich, Germany)

Two topics are discussed in the following. First, red cell lysis and membrane transfer are primarily caused by cavitation since they are suppressed by excess hydrostatic pressure in the exposure vessel. It was additionally proposed that shock waves destroyed red blood cells directly at excess pressure above 10 MPa when cavitation was absent. When this was re-examined with a different pressure chamber with 10 and 20 MPa excess pressure, there was no increased red cell lysis and no direct action of shock waves was found. Second, it was long thought that the transfer of fluorescein dextran into cells was a good method to co-transfer another molecule. Dextran and the other substance had to be dissolved well in a defined molecular ratio and cells with a defined number of dextran molecules recovered by flow sorting contained also a defined number of the other molecule. The attempt to apply this approach for various numbers of molecules of a ribosome inactivating protein revealed, however, an inconclusive result.

10:05–10:25 Break

10:25

1aBB9. The disruption of tissue structure using high intensity pulsed ultrasound. J. Brian Fowlkes (Dept. of Radiol., Univ. of Michigan, Kresge III R3315, Ann Arbor, MI 48109-0553, fowlkes@umich.edu), Jessica E. Parsons, Zhen Xu, Michol Cooper, Binh C. Tran, Timothy L. Hall, William W. Roberts, and Charles A. Cain (Univ. of Michigan, Ann Arbor, MI 48109-0330)

Recent investigations of pulsed ultrasound at high acoustic intensities have revealed a regime in which significant breakdown of tissue structure can be achieved. This therapeutic modality, which might be termed histotripsy, is dependent on the presence of highly active cavitation evidenced by significant temporal fluctuations in acoustic backscatter. In the presence of tissue interfaces, erosion can result yielding, for example, well-defined perforations potentially useful in creating temporary shunts for the treatment of hypoplastic left heart syndrome. When applied in bulk tissue, the process results in a near emulsification with little structural integrity remaining or chance of cellular survival. In each case, the process is dependent on acoustic parameters of the field to not only produce damage for a given pulse but also to sustain the cavitation nuclei population for subsequent pulses. Fluctuations in acoustic backscatter indicate both initiation and extinction of the appropriate cavitation activity during application of therapeutic ultrasound, which leads to a potential feedback mechanism to minimize acoustic exposure. This presentation will discuss the observed tissue damage as affected by acoustic parameters and the ability to monitor the presence of cavitation activity expected to be responsible for these effects. [Work supported by NIH grants RO1 RR14450.]

Contributed Papers

10:40

1aBB10. Observation of cavitation during shock wave lithotripsy. Michael R. Bailey, Lawrence A. Crum (Ctr. for Industrial and Medical Ultrasound, Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105), Yuri A. Pishchalnikov, James A. McAteer, Irina V. Pishchalnikova, Andrew P. Evan (Indiana Univ. School of Medicine, Indianapolis, IN 46202-5120), Oleg A. Sapozhnikov (M.V. Lomonosov Moscow State Univ., Moscow, 119992, Russia), and Robin O. Cleveland (Boston Univ., Boston, MA 02215)

A system was built to detect cavitation in pig kidney during shock wave lithotripsy (SWL) with a Dornier HM3 lithotripter. Active detection, using echo on B-mode ultrasound, and passive cavitation detection (PCD), using coincident signals on bifocal, orthogonal receivers, were equally sensitive and were used to interrogate the renal collecting system (urine) and the kidney parenchyma (tissue). Cavitation was detected in urine immediately upon SW administration in urine or urine plus X-ray contrast agent, but in tissue, cavitation required hundreds of SWs to initiate. Localization of cavitation was confirmed by fluoroscopy, sonography, and by thermally marking the kidney using the PCD receivers as high intensity focused ultrasound sources. Cavitation collapse times in tissue and native urine were about the same but less than in urine after injection of X-ray contrast agent. Cavitation, especially in the urine space, was observed to evolve from a sparse field to a dense field with strong acoustic collapse emissions to a very dense field that no longer produced detectable collapse. The finding that cavitation occurs in kidney tissue is a critical step toward determining the mechanisms of tissue injury in SWL. [Work sup

ported by NIH (DK43881, DK55674, FIRCA), ONRIFO, CRDF and NS-BRI SMS00203.]

10:55

1aBB11. Detecting cavitation *in vivo* from shock-wave therapy devices. Thomas J. Matula, Jinfei Yu, and Michael R. Bailey (Ctr. for Industrial and Medical Ultrasound, Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

Extracorporeal shock-wave therapy (ESWT) has been used as a treatment for plantar fasciitis, lateral epicondylitis, shoulder tendonitis, non-unions, and other indications where conservative treatments have been unsuccessful. However, in many areas, the efficacy of SW treatment has not been well established, and the mechanism of action, particularly the role of cavitation, is not well understood. Research indicates cavitation plays an important role in other ultrasound therapies, such as lithotripsy and focused ultrasound surgery, and in some instances, cavitation has been used as a means to monitor or detect a biological effect. Although ESWT can generate cavitation easily *in vitro*, it is unknown whether or not cavitation is a significant factor *in vivo*. The purpose of this investigation is to use diagnostic ultrasound to detect and monitor cavitation generated by ESWT devices *in vivo*. Diagnostic images are collected at various times during and after treatment. The images are then post-processed with image-processing algorithms to enhance the contrast between bubbles and surrounding tissue. The ultimate goal of this research is to utilize cavitation as a means for optimizing shock wave parameters such as amplitude and pulse repetition frequency. [Work supported by APL internal funds and NIH DK43881 and DK55674.]

1aBB12. Secondary shock wave emissions from cavitation in lithotripsy. Parag V. Chitnis and Robin O. Cleveland (Aerosp. and Mech. Eng. Dept., Boston Univ., 110 Cummington St., Boston, MA 02215)

We investigate the role of secondary shock waves (SSWs) generated by cavitation in lithotripsy. Acoustic pressure was measured with a fiber optic probe hydrophone and cavitation using a dual passive cavitation detector (PCD) consisting of two confocal transducers. An artificial stone (~7 mm diameter and ~9 mm length) was placed at the focus of an electrohydraulic lithotripter. The fiber was inserted through a hole drilled through the stone so that the tip was at the proximal surface. SSWs were

identified by matching the time of arrival to that of the inertial collapse signature acquired by the PCD. Measurements of SSWs were obtained for 50% of SWs fired at 20 kV and 1 Hz. The peak positive pressure for the SSW was $p^+ = 33.7 \pm 14.8$ MPa, which was comparable to the pressure induced by the incident SW ($p^+ = 42.6 \pm 6$ MPa). The peak pressure in water was $p^+ = 23.2 \pm 4.4$ MPa. The PCD also recorded acoustic emissions from forced collapse of pre-existing bubbles caused by the incident SW. We propose that both the reflection from the semi-rigid stone boundary and SSW from the forced collapse contribute to the observed increase in the peak pressure of the incident SW in presence of a stone. [Work supported by NIH.]

MONDAY MORNING, 16 MAY 2005

REGENCY C, 8:00 TO 11:35 A.M.

Session 1aPP

Psychological and Physiological Acoustics and Animal Bioacoustics: Size Information in Speech and Animal Calls

Roy D. Patterson, Cochair

Physiology Dept., Univ. of Cambridge, Downing St., Cambridge CB2 3EG, United Kingdom

Toshio Irino, Cochair

Wakayama Univ., Systems Engineering, 930 Sakaedami, Wakayama 640-8510, Japan

Chair's Introduction—8:00

Invited Papers

8:05

1aPP1. Psychoacoustic evaluation of a low-parameter modal model for synthesizing impact sounds. Robert A. Lutfi, Eileen Storm, Joshua M. Alexander (Dept. of Communicative Disord. and Waisman Ctr., Univ. of Wisconsin, Madison, WI 53706, ralutfi@wisc.edu), and Eunmi Oh (Samsung AIT, Suwon, Korea 440-600)

Three experiments were conducted to test the viability of a low-parameter modal model for synthesizing impact sounds to be used in commercial and psychoacoustic research applications. The model was constrained to have 4 physically-based parameters dictating the amplitude, frequency and decay of modes. The values of these parameters were selected by ear roughly to match the recordings of 10 different resonant objects suspended by hand and struck with different mallets. In Exp. 1, neither 35 professional musicians nor 187 college undergraduates could identify which of the 2 matched sounds was the real recording with better than chance accuracy, though significantly better than chance performance was obtained when modal parameters were selected without the previously imposed physical constraints. In Exp. 2, the undergraduates identified the source corresponding to the recorded and synthesized sounds with the same level of accuracy and largely the same pattern of errors. Finally, Exp. 3 showed highly-practiced listeners to be largely insensitive to changes in the acoustic waveform resulting from an increase in the number of free-parameters used in the modal model beyond 3. The results suggest that low-parameter, modal models might be meaningfully exploited in many commercial and research applications involving human perception of impact sounds.

8:20

1aPP2. Reliable but weak voice-formant cues to body size in men but not women. Drew Rendall, John R. Vokey, Christie Nemeth, and Christina Ney (Dept of Psych., Univ. of Lethbridge, Lethbridge, AB, Canada, d.rendall@uleth.ca)

Whether voice formants provide reliable cues to adult body size has been contested recently for some animals and humans and the outcome bears critically on theories of social competition and mate choice, language origins, and speaker normalization. We report two experiments to test listeners' ability to assess speaker body size. In Experiment 1, listeners heard paired comparisons of the same short phrase spoken by two adults of the same sex paired randomly with respect to height and indicated which was larger. Both sexes (M=20; F=22) showed an equal but modest ability to identify the larger male (mean correct=58.5%; $T=31.5$, $P<0.001$) that correlated with the magnitude of their height difference but could not pick the larger female (mean correct=52.0%; $T=1.05$, $P=0.305$) regardless of the height difference. Experiment 2 used single word comparisons, focused only on male voices, and controlled F_0 while manipulating F_1-F_4 between speakers. When F_0 was equal but F_1-F_4 predicted the height difference between speakers, both sexes (M=12; F=18) correctly chose the taller male (80%). When F_1-F_4 values of the shorter male were reduced below those of the taller male (or vice versa), subjects shifted to pick the shorter male as being larger.

8:35

1aPP3. Formant frequencies as indicators of body size in red deer roars. David Reby, Karen McComb, Chris Darwin (Dept. of Psych., Univ. of Sussex, Brighton, BN2 9QH, UK), and W. Tecumseh Fitch (Univ. of St. Andrews, St. Andrews, KY16 9JU, UK)

Recent studies of animal vocal communication have emphasized the potential for vocal tract resonances to encode information on the size of callers and the need for receivers to attend to this information, in particular in the context of intra-sexual competition and inter-sexual mate choice. Our recent work on red deer roaring, a classical example of a sexual communication signal, is reviewed here. A combination of anatomical analyses of the vocal apparatus, acoustical analyses, and playback experiments using re-synthesized calls has enabled us to show that: (i) red deer and fallow deer males have a descended and mobile larynx, an anatomical innovation that was previously believed to be uniquely human and that enables callers to modulate their formants during vocalizing; (ii) minimum formant frequencies provide an honest indication of body size in red deer roars and (iii) stags use rivals' minimum formant frequencies in assessment during male-male contests, and adjust the formants of their own replies in relation to what they hear.

8:50

1aPP4. Extracting a carrier-independent version of the syllabic message: The principles. Roy D. Patterson, Thomas C. Walters (Ctr. for the Neural Basis of Hearing, Physio. Dept., Univ. of Cambridge, Downing St., Cambridge, CB2 3EG, UK, rdp1@cam.ac.uk), and Toshio Irino (Wakayama Univ., 930 Sakaedani Wakayama 640-8510, Japan)

At the heart of each syllable of speech is a vowel; the wave consists of a stream of glottal pulses, each with a resonance attached. The vowel contains three important components of the information in the larger communication: the glottal pulse rate (the pitch), the resonance shape (the message), and the resonance scale (the vocal tract length). Recent experiments on the perception of vowels show that variability in glottal pulse rate and vocal tract length has surprisingly little effect on the humans ability to recognise the vowel or discriminate speaker size, despite the variability it imparts to the spectra of these sounds. We appear to have an automatic normalization process to scale vowels and extract the message independent of the carrier. Many animal calls are like syllables in form and duration, and normalization is essential here as well if animals are to correctly identify the species of the sender and not be confused by changes in pulse rate and resonance scale that simply indicate a size difference. This talk describes how neural firing patterns produced by vowels and animal calls could be normalized to produce a carrier independent version of the message of the syllable. [Work supported by UK MRC.]

9:05

1aPP5. The stabilized, wavelet-Mellin transform for analyzing the size and shape information of vocalized sounds. Toshio Irino (Faculty of Systems Eng., Wakayama Univ., 930 Sakaedani, Wakayama 640-8510, Japan) and Roy Patterson (Univ. of Cambridge, Cambridge, CB2 3EG, UK)

We hear vowels produced by men, women, and children as approximately the same although there is considerable variability in glottal pulse rate and vocal tract length. At the same time, we can identify the speaker group. Recent experiments show that it is possible to identify vowels even when the glottal pulse rate and vocal tract length are condensed or expanded beyond the range of natural vocalization. This suggests that the auditory system has an automatic process to segregate information about shape and size of the vocal tract. Recently we proposed that the auditory system uses some form of Stabilized, Wavelet-Mellin Transform (SWMT) to analyze scale information in bio-acoustic sounds as a general framework for auditory processing from cochlea to cortex. This talk explains the theoretical background of the model and how the vocal information is normalized in the representation. [Work supported by GASR(B)(2) No. 15300061, JSPS.]

9:20

1aPP6. Manipulating the pulse rate and resonance scale in speech and animal calls. Hideki Kawahara (930 Sakaedani, Wakayama, 640-8510, Japan)

A large proportion of the sound in speech and animal calls is voiced, and thus, periodic in nature. The sounds are generated by repetitive pulsed stimulations and each pulse produces a resonant response. This repetitive structure in the sounds can be interpreted as a time-frequency sampling process that provides a stream of information about the size, shape, and structure of the resonators in the vocal tract. This perspective has enabled us to develop a system, referred to as STRAIGHT, that can analyze, manipulate, and resynthesize vocal sounds. It is based on an extended pitch synchronous spectral estimation algorithm that recovers the underlying smooth time-frequency representation representing physical information on resonators. Ideally, this process removes stimulation related structure from the time-frequency representation, and thus, it should follow the same scaling laws as the physical dimensions of the resonating body. One problem is that speech sometimes contains multiple stimulations within one pitch period. This type of stimulation introduces a spectral deformation that has the same scaling laws as for the fundamental frequency (reciprocal of the repetition rate). The effects of this dual scaling and the problems of joint normalization will be discussed. [Work supported by MEXT Japan.]

9:50

1aPP7. Relationship between fundamental and formant frequencies in speech perception. Peter F. Assmann (School of Behavioral and Brain Sci., Univ. of Texas at Dallas, Box 830688, Richardson, TX 75083), and Terrance M. Nearey (Univ. of Alberta, Edmonton, Alberta, Canada T6E 2G2)

In natural speech, there is a moderate correlation between fundamental frequency (F_0) and formant frequencies, associated with differences in larynx and vocal tract size across talkers. We have manipulated these properties using the Kawahara's STRAIGHT vocoder to determine their contribution to sentence intelligibility, vowel identification accuracy, perceived gender, and naturalness judgments by human listeners. The results of these experiments, together with predictions from a pattern recognition model, suggest that frequency-shifted speech is more intelligible and perceived as more natural when the natural co-variation of F_0 and formant frequency is preserved, even when frequency shifts exceed the range found in human speech.

10:05

1aPP8. Perception of speaker size and sex of vowel sounds. David R. R. Smith and Roy D. Patterson (CNBH, Dept. of Physiol., Cambridge Univ., Downing St., Cambridge CB2 3EG, UK, david.smith@mrc-cbu.cam.ac.uk)

Glottal-pulse rate (GPR) and vocal-tract length (VTL) are both related to speaker size and sex—however, it is unclear how they interact to determine our perception of speaker size and sex. Experiments were designed to measure the relative contribution of GPR and VTL to judgements of speaker size and sex. Vowels were scaled to represent people with different GPRs and VTLs, including many well beyond the normal population values. In a single interval, two response rating paradigm, listeners judged the size (using a 7-point scale) and sex/age of the speaker (man, woman, boy, or girl) of these scaled vowels. Results from the size-rating experiments show that VTL has a much greater influence upon judgements of speaker size than GPR. Results from the sex-categorization experiments show that judgements of speaker sex are influenced about equally by GPR and VTL for vowels with normal GPR and VTL values. For abnormal combinations of GPR and VTL, where low GPRs are combined with short VTLs, VTL has more influence than GPR in sex judgements. [Work supported by the UK MRC (G9901257) and the German Volkswagen Foundation (VWF 1/79 783).]

10:20

1aPP9. Identification of “size-modulated” vowels sequences: Effects of modulation periods and speaking rates. Minoru Tsuzaki (Kyoto City Univ. of Arts, 13-6 Kutsukake-cho, Oe, Nishikyo-ku, Kyoto, 610-1197 Japan, minoru.tsuzaki@kcuu.ac.jp), Toshio Irino (Wakayama Univ., Wakayama, Wakayama 640-8510, Japan), and Roy Patterson (Univ. of Cambridge, Cambridge, CB2 3EG, UK)

We investigated the temporal dynamics of auditory normalization and size perception by measuring vowel recognition performance using sequences of vowels in which vocal tract length was modulated during the sequence. The modulation of speaker size was achieved by scaling the frequency axis of the transfer function of vocal tract. The temporal modulation pattern was sinusoidal with a period of 16, 32, 64, 128, 256, 512, 1024, 2048, or 4096 ms. Listeners identified sequences of six vowels from four response alternatives. Although the listeners had no experience with size-modulated speech, the percentage of correct responses was never less than 90% for any modulation period. This suggests that the auditory system has an automatic size-normalizing mechanism which does not require training. The listeners had most difficulty with the 256-ms modulation period, independent of the speaking rate of the sequence. This might indicate a limitation of the processing speed for size-normalization. A simulation using Mellin Images did not reveal any obvious reason for the dip in performance with the 250 ms period, which suggests that the limitation is not in the image construction stage. [Work supported by GASR(A)(2) No. 16200016, JSPS.]

10:35

1aPP10. The perception of size in musical instrument sounds. Ralph van Dinter and Roy D. Patterson (CNBH, Dept. of Physiol., Univ. of Cambridge, Downing St., Cambridge CB2 3EG, UK, ralph.van-dinter@mrc-cbu.cam.ac.uk)

There is size information in natural sounds. For example, as humans grow in height their vocal tract increases in length, and this produces a predictable decrease in formant frequency. Recent studies have shown that listeners can judge the relative size of two individuals with considerable precision, and they can recognize vowels scaled well beyond the range normally experienced. This paper reports two psychophysical studies designed to extend the research to musical instrument sounds. The first showed that listeners can discriminate the relative size of instruments reliably, although not quite as well as for voices. The second showed that listeners can recognize instrument sounds scaled in size and pitch well beyond the range of normal experience. The research supports the hypothesis that the auditory system applies some kind of active normalization to all input sounds. [Work supported by the U.K. Medical Research Council (G9901257, G9900362), and ONRIFO (Grant N00014-03-1-1023).]

11:05

1aPP11. The neural processing of musical instrument size information in the brain investigated by magnetoencephalography. Andre Rupp (Section of Biomagnetism, Dept. Neurology, Univ. of Heidelberg, Im Neuenheimer Feld 400, 69120 Heidelberg, Germany), Ralph van Dinther, and Roy D. Patterson (Cambridge Univ., UK)

The specific cortical representation of size was investigated by recording auditory evoked fields (AEFs) elicited by changes of instrument size and pitch. In Experiment 1, a French horn and one scaled to double the size played a three note melody around $F3$ or its octave, $F4$. Many copies of these four melodies were played in random order and the AEF was measured continuously. A similar procedure was applied to saxophone sounds in a separate run. In Experiment 2, the size and type of instrument (French horn and saxophone) were varied without changing the octave. AEFs were recorded in five subjects using magnetoencephalography and evaluated by spatio-temporal source analysis with one equivalent dipole in each hemisphere. The morphology of the source waveforms revealed that each note within the melody elicits a well-defined $P1-N1-P2$ AEF-complex with adaptation for the 2nd and 3rd note. At the transition of size, pitch, or both, a larger AEF-complex was evoked. However, size changes elicited a stronger $N1$ than pitch changes. Furthermore, this size-related $N1$ enhancement was larger for French horn than saxophone. The results indicate that the $N1$ plays an important role in the specific representation of instrument size.

11:20

1aPP12. Voices of athletes reveal only modest acoustic correlates of stature. Michael J. Owren and John D. Anderson IV (Dept. of Psych., Cornell Univ., Ithaca, NY 14853)

Recent studies of acoustic cues to body-size in nonhuman primate and human vocalizations have produced results varying from very strong relationships between formant frequencies and length/weight in rhesus monkeys to weak correlations between formants and stature in humans. The current work attempted to address these discrepancies by compiling a database of naturally occurring speech with a large number of vocalizers of maximally varying size. To that end, fundamental frequency ($F0$) and formant frequencies were measured in both running speech and filled pauses (i.e., “ah” and “um”) produced by male athletes during televised same-day interviews. Multiple-regression analysis of data from 100 male athletes showed that these acoustic measures accounted for at most 17% of variance in height over a 37-cm range. Analyses of filled speech pauses produced by a subset of 48 athletes could account for up to 36%. These outcomes fall within the range of previously reported outcomes, indicating that while speech acoustics are correlated with body-size in human adult males, the cues provided are quite modest.

MONDAY MORNING, 16 MAY 2005

REGENCY D, 8:30 TO 11:35 A.M.

Session 1aSC

Speech Communication and Psychological and Physiological Acoustics: Communication Abilities of Congenitally Deaf Children: From Behavior to Physiology From Psychophysics to Hair Cell Regeneration

Mario A. Svirsky, Cochair

Indiana Univ., School of Medicine, ENT, 699 West Dr., Indianapolis, IN 46202

Ruth Y. Litovsky, Cochair

Univ. of Wisconsin-Madison, Waisman Center, 1500 Highland Ave., Madison, WI 53705

Chair's Introduction—8:30

8:35

1aSC1. Language acquisition after cochlear implantation of congenitally deaf children: Effect of age at implantation. Mario Svirsky and Rachael Holt (Indiana Univ. School of Med., 699 West Dr., RR-044, Indianapolis, IN 46202)

Evidence shows that early implantation of congenitally deaf children is beneficial. However, infants as young as 6 months of age have started to receive cochlear implants (CIs) in the USA. Such early implantation may be associated with higher risks, including anesthetic risk as well as the increased possibility of a false positive in the diagnosis of profound deafness. On the other hand, delaying implantation may be associated with the risk of missing windows of opportunity or sensitive periods for the development of communication skills. In this study, speech perception and language skills in children who received CIs in the first, second, third, or fourth year of life were compared. Participants were tested at regular 6-month intervals after implantation. The effects of several potential confounds were considered. In general, children implanted earlier outperformed those implanted later, with one exception: infants implanted at 6–12 months showed similar outcomes to children implanted at 12–24 months, at least through 2 to 2-1/2 years of age. This preliminary result may be associated with the difficulty of choosing appropriate stimulation parameters for infants, and its potential influence on the quality of the stimulation patterns delivered by the CI.

9:00

1aSC2. Phonological systems of pediatric cochlear implant users: The acquisition of voicing. Steven B. Chin (Dept. of Otolaryngol., Indiana Univ. School of Medicine, 699 West Dr., RR044, Indianapolis, IN 46202-5119, schin@iupui.edu), Eric N. Oglesbee, Andrew K. Kirk (Indiana Univ., Bloomington, IN), and Joseph E. Krug (Indiana Univ. School of Medicine, Indianapolis, IN)

Although cochlear implants are primarily auditory prostheses, they have also demonstrated their usefulness as aids to speech production and the acquisition of spoken language in children. This presentation reports on research currently being conducted at the Indiana University Medical Center on the development of phonological systems by children with five or more years of cochlear implant use in English-speaking environments. Characteristics of the feature [voice] will be examined in children with cochlear implants and in two comparison groups: adults with normal hearing and children with normal hearing. Specific aspects of voicing to be discussed include characteristic error patterns, phonetic implementation of the voicing contrast, and phonetic implementation of neutralization of the voicing contrast. Much of the evidence obtained thus far indicates that voicing acquisition in children with cochlear implants is not radically different from that of children with normal hearing. Many differences between the systems of children with cochlear implants and the ambient system thus appear to reflect the children's age as much as their hearing status. [Work supported by grants from the National Institutes of Health to Indiana University: R01DC005594 and R03DC003852.]

9:25

1aSC3. Source reconstruction of sensory and cognitive evoked potentials. Curtis W. Ponton (Compumedics Neuroscan, 7850 Paseo del Norte, El Paso, TX 79912)

Cortical activity reflected in auditory-evoked potentials (AEPs) is often evaluated using only a small subset of all the recorded data. Conclusions based on this approach can be misleading, with no utility in identifying the underlying neural generators of the scalp recorded activity. Techniques that make use of all of the AEP data range from relatively simple methods, such as global field power, to more advanced approaches including independent components analysis (ICA), dipole modeling, and current density reconstruction. The objective of this presentation is to describe analysis of the component-structure of AEPs using these advanced techniques. Results of the ICA analysis will be used to generate spatial filters characterizing the specific scalp distribution associated with each ICA pattern. The cortical origin of each ICA scalp distribution will then be determined using current density reconstruction. The analysis will be applied to standard AEPs as well as the mismatch negativity in normal hearing and cochlear implant users. The results will demonstrate the unique suitability of neuroimaging based on AEPs for understanding the effects of cochlear implant use on cortical activity associated with cognitive processing in children and in adults.

9:50–10:05 Break

10:05

1aSC4. Auditory plasticity in deaf children with bilateral cochlear implants. Ruth Litovsky (Waisman Ctr., Univ. of Wisconsin, Madison, WI 53705, litovsky@waisman.wisc.edu)

Human children with cochlear implants represent a unique population of individuals who have undergone variable amounts of auditory deprivation prior to being able to hear. Even more unique are children who received bilateral cochlear implants (BICIs), in sequential surgical procedures, several years apart. Auditory deprivation in these individuals consists of a two-stage process, whereby complete deafness is experienced initially, followed by deafness in one ear. We studied the effects of post-implant experience on the ability of deaf children to localize sounds and to understand speech in noise. These are two of the most important functions that are known to depend on binaural hearing. Children were tested at time intervals ranging from 3-months to 24-months following implantation of the second ear, while listening with either implant alone or bilaterally. Our findings suggest that the period during which plasticity occurs in human binaural system is protracted, extending into middle-to-late childhood. The rate at which benefits from bilateral hearing abilities are attained following deprivation is faster for speech intelligibility in noise compared with sound localization. Finally, the age at which the second implant was received may play an important role in the acquisition of binaural abilities. [Work supported by NIH-NIDCD.]

10:30

1aSC5. Learning on auditory discrimination tasks in normal-hearing listeners: Implications for hearing rehabilitation. Beverly A. Wright (Dept. of Commun. Sci. and Disord., 2240 Campus Dr., Northwestern Univ., Evanston, IL 60208-3550, b-wright@northwestern.edu)

Hearing rehabilitation extends beyond simply fitting a hearing aid or cochlear implant. To improve the benefit of these devices, it must be established which auditory abilities can be improved with training. Toward this end, learning in normal-hearing listeners was examined on five auditory discrimination tasks: frequency, intensity, interaural-time-difference (ITD), interaural-level-difference (ILD), and duration. Because the same training regimen was used throughout, any differences in the learning patterns across these trained discriminations likely reflect differences in the plasticity of the underlying mechanisms, at least for that regimen. The influence of training was assessed by comparing the improvements in discrimination threshold on trained and untrained conditions between listeners who were given multiple-hour practice on a single discrimination condition and those who were not. Learning on the five tasks followed one of two general patterns. For ITD and intensity discrimination, multiple-hour practice did not lead to greater learning than that seen in untrained listeners. In contrast, for ILD, duration, and frequency discrimination, such practice yielded greater learning, but only on a subset of conditions. The differences in the plasticity across these auditory tasks in normal-hearing listeners imply that cochlear-implant users may benefit more from training on some tasks than others. [Work supported by NIH.]

1aSC6. Hair cell regeneration: Look to the future. Edwin W. Rubel (Virginia Merrill Bloedel Hearing Res. Ctr., Univ. of Washington, M.S. 357923, Seattle, WA 98195)

Less than 2 decades ago it was discovered that birds can regenerate hair cells in the auditory and vestibular parts of the inner ear after the native hair cells are destroyed by exposure to excessive noise or by mechanical trauma of aminoglycoside antibiotics. This discovery issued in a new era of hearing research—it suggested that some day it may be possible to actually restore hearing in people with congenital or acquired hearing loss due to the degeneration of sensory cells or supporting cells in the inner ear. Fifteen years is a very short time in the history of science. Consider the fact that we have actively sought chemical treatments to prevent or cure cancers for well over a half century and the “war on Cancer,” resulted in enormous public and private support. Progress has been great, and some forms of cancer can be treated with great success, but the overall 5-year survival rates have only risen from about 50% to 63%. Progress will continue and many more forms of cancer will be cured and prevented during the next half century. Similarly, during the first 15 years of hair cell regeneration research enormous progress has been made, and we now know that postnatal mammalian ears have the capacity to produce new hair cells. We are indeed a long way from restoring hearing through hair cell regeneration, but the future is pretty clear. I will review the progress of this field with an eye toward the future and what it means for treatments of today. In particular, I will address the potential cost versus benefits of bilateral implantation when applied to babies and young children.

Contributed Paper

11:20

1aSC7. Preliminary comparison of infants speech with and without hearing loss. Richard S. McGowan (CReSS LLC, 1 Seaborn Pl., Lexington, MA 02420), Susan Nittrouer (Utah State Univ., Logan, UT 84322), and Karen Chenausky (Speech Technol. & Appl. Res., Bedford, MA 01730)

The speech of ten children with hearing loss and ten children without hearing loss aged 12 months is examined. All the children with hearing loss were identified before six months of age, and all have parents who

wish them to become oral communicators. The data are from twenty minute sessions with the caregiver and child, with their normal prostheses in place, in semi-structured settings. These data are part of a larger test battery applied to both caregiver and child that is part of a project comparing the development of children with hearing loss to those without hearing loss, known as the Early Development of Children with Hearing Loss. The speech comparisons are in terms of number of utterances, syllable shapes, and segment type. A subset of the data was given a detailed acoustic analysis, including formant frequencies and voice quality measures. [Work supported by NIDCD R01 006237 to Susan Nittrouer.]

MONDAY AFTERNOON, 16 MAY 2005

REGENCY E, 1:00 TO 5:20 P.M.

Session 1pAA

Architectural Acoustics: Architectural Acoustics and the Green Building Movement

Ralph T. Muehleisen, Cochair

Illinois Inst. of Technology, Civil and Architectural Engineering, 3201 South Dearborn, Chicago, IL 60616

Murray R. Hodgson, Cochair

Univ. of British Columbia, School of Occupational and Environmental Hygiene, 2206 East Mall, Vancouver, BC V6T 1Z3, Canada

Chair's Introduction—1:00

Invited Papers

1:05

1pAA1. Assessing the future of green building. Raymond J. Cole (UBC School of Architecture, 6333 Memorial Rd., Vancouver, BC V6T 1Z2, Canada, raycole@arch.ubc.ca)

As the realities of resource depletion and global environmental degradation become more evident, we can anticipate a maturing and strengthening of the public's concern and knowledge on environmental issues. This will translate into an expectation for greater environmental responsibility and, as with other sectors, the building industry will be increasingly scrutinized for its environmental actions. The adoption of environmental strategies has been accelerated by the emergence of building environmental assessment methods that have provided both a definition and common language for green buildings as well as a means of communicating performance improvements. Whereas the current focus is on “green” design—reducing or mitigating the environmental consequences of buildings—the future concerns will embrace mitigation, adaptation to the new conditions and restoring previous adversely impacted regions and human settlements. This presentation will provide an overview of the evolution of green building practices to set a context for understanding emerging issues in building acoustics. Since the adoption of green building practices is a function of the context that shapes political and public priorities, the presentation compares and contrasts several short and long-term scenarios some certain, others more speculative and their direct and indirect consequences for environmental progress building design.

1:35

1pAA2. Integrated building design. Jennifer Sanguinetti (Keen Eng., 116-930 West First St., North Vancouver, BC, Canada V7P 3N4, Jennifer.Sanguinetti@keen.ca)

For many years, building design has been a very linear process with owners speaking to architects who then design building shells that they pass along to sub-consultants who must fit their systems into the allotted spaces. While this process has some advantages, it provides little opportunity to optimize systems based on such factors as energy use or occupant comfort. This presentation will focus on the evolution and implications of integrated building design, a method that has provided greater opportunities for interaction between design disciplines and with building users early on in the design process. Integration has resulted in buildings that are more sustainable than typical buildings and that can respond better to the needs of the owner and users. Examples of the application of the process and the resulting buildings will be presented from the view of a design engineer with experience of both processes. Specifically, the potential contribution of an acoustical consultant in the integrated process will be explored.

2:05

1pAA3. Post-occupancy evaluation and acoustics in green buildings. Rosamund Hyde (Keen Eng., 116-930 West First St., North Vancouver, BC, Canada V7P 3N4, Rosie.Hyde@keen.ca)

Post-occupancy evaluation is a process for discerning whether completed buildings are delivering to their occupants and owners the benefits which were set as goals in the design process. Both resource consumption and occupant satisfaction are assessed. Evaluations of seven green buildings in British Columbia have been carried out, including buildings with a range of occupancies and climates. These studies have highlighted acoustics as an area where occupant satisfaction could be improved. This presentation will address three questions. First, what have we learned about acoustics from post-occupancy evaluation of green buildings? Results of the seven post-occupancy evaluations of green buildings will be discussed. Second, what questions do designers have about decisions related to acoustics in green buildings? Third, how can we improve post-occupancy evaluation processes to better address those questions?

2:35

1pAA4. Green buildings: Implications for acousticians. Michael R. Noble (BKL Consultants Ltd., 308-1200 Lynn Valley Rd., North Vancouver, BC, Canada V7J 2A2, noble@bkla.com)

This presentation will deal with the practical implications of green design protocols of the US Green Building Council on interior acoustics of buildings. Three areas of particular consequence to acousticians will be discussed. Ventilation Systems: reduced energy consumption goals dictate reliance on natural cooling and ventilation using ambient air when possible. The consequent large openings in the building envelope to bring fresh air into rooms, and similar sized openings to transfer the mixed air out, can severely compromise the noise isolation of the rooms concerned. Radiant Cooling: the heavy concrete floors of buildings can be used as a thermal flywheel to lessen the cooling load, which forces the concrete ceilings to be exposed to the occupied rooms for heat transfer, and strictly limits the application of acoustical absorption on the ceilings. This challenges the room acoustics design. Green Materials: the LEED protocols require the elimination of potentially harmful finishes, including fibrous materials which may impact air quality or contribute to health problems. Since the backbone of sound absorption is glass and mineral fibres, this further challenges provision of superior room acoustics. Examples and commentary will be provided based on current and recent projects.

3:05–3:20 Break

Contributed Papers

3:20

1pAA5. A comparative study of varying fan noise mitigation techniques in relation to sustainable design goals. Cornelis Overweg (Acentech, 1429 E. Thousand Oaks Blvd., Ste. 200, Thousand Oaks, CA 91362) and Jeff L. Fullerton (Acentech Inc., Cambridge, MA 02138)

Green building design promotes effective use of materials and energy, improved indoor environmental quality (IEQ), and enhanced occupant comfort. These “green” goals can occasionally conflict with common acoustical approaches used for fan noise control. A design striving for low noise levels from the ventilation system to benefit occupant comfort can inadvertently introduce elements that are contradictory to other green building objectives. For example, typical fan noise control devices introduce higher energy consumption or less beneficial indoor environmental quality. This paper discusses the acoustical, mechanical, environmental, and relative cost impacts of various fan noise control techniques.

3:35

1pAA6. Acoustical case studies of three green buildings. Gary Siebein (Univ. of Florida School of Architecture, P.O. Box 115702, Gainesville, FL 32607), Robert Lilkendey, and Stephen Skorski (Siebein Assoc., Inc., Consultants in Architectural Acoust., Gainesville, FL 32607)

Case studies of 3 green buildings with LEED certifications that required extensive acoustical retrofit work to become satisfactory work environments for their intended user groups will be used to define areas where green building design concepts and acoustical design concepts require reconciliation. Case study 1 is an office and conference center for a city environmental education agency. Large open spaces intended to collect daylight through clerestory windows provided large, reverberant volumes with few acoustic finishes that rendered them unsuitable as open office space and a conference room/auditorium. Case Study 2 describes one of the first gold LEED buildings in the southeast whose primary design concepts were so narrowly focused on thermal and lighting issues that they often worked directly against basic acoustical requirements re-

sulting in sound levels of NC 50–55 in classrooms and faculty offices, crosstalk between classrooms and poor room acoustics. Case study 3 is an environmental education and conference center with open public areas, very high ceilings, and all reflective surfaces made from wood and other environmentally friendly materials that result in excessive loudness when the building is used by the numbers of people which it was intended to serve.

3:50

1pAA7. An acoustic window for sustainable buildings. Jian Kang, Martin Brocklesby, Zheming Li (School of Architecture, Univ. of Sheffield, Western Bank, Sheffield S10 2TN, UK, j.kang@sheffield.ac.uk), and David J. Oldham (Univ. of Liverpool, L69 3BX, UK)

Encouraging the use of natural ventilation is an important tendency in the green building movement, but opening windows can often cause noise problems. This research develops a window system which allows natural ventilation while reducing noise transmission. The core idea is to create a ventilation path by staggering two layers of glass and using micro-perforated absorbers (MPA) along the path created to reduce noise. The MPA are made from transparent materials so that daylighting is relatively unaffected. Starting with a brief introduction of the MPA theory and its application in ducts, the paper presents a series of numerical simulations using finite element method based software FEMLAB, and experiment results measured between a semi-anechoic chamber and a reverberation chamber. Performance in acoustics, ventilation and daylighting are all taken into account. A basic window configuration is first considered, studying the effectiveness of various window parameters. A number of

strategic designs are then examined, including external hoods and louvers in the sound path. There is generally a good agreement between simulation and measurement, and the noise reduction can be as good as a single glazing, with air movement to achieve occupant comfort, rather than just for minimum air exchange. [Work supported by EPSRC.]

4:05

1pAA8. Modeling the acoustical and airflow performance of natural ventilation inlet and outlet units. David J. Oldham (School of Architecture, Univ. of Liverpool, Liverpool, L69 3BX, UK, djoldham@liv.ac.uk), Jian Kang, and Martin Brocklesby (Univ. of Sheffield, Western Bank, Sheffield S10 2TN, UK)

One aspect of the trend towards designing green buildings has been the increasing use of natural ventilation for buildings which otherwise might have required mechanical ventilation or even full air conditioning. However, the pressure differentials available to drive the natural ventilation process are small and hence relatively large inlets and outlets with low resistance to flow are required. These apertures constitute significant acoustic weak points on building facades and hence need to be treated to reduce noise ingress. Although there are a number of natural ventilation units available they have frequently been designed from the application of simple principles without any attempt to optimise both their airflow and acoustical performance. In this paper the results of a series of computer modeling exercises are described using acoustic FEM and BEM plus Computational Fluid Dynamics (CFD) which seeks to establish recommendations for the optimum design of natural ventilation inlet and outlet devices for both acoustical and airflow performance.

4:20–5:20

Panel Discussion

MONDAY AFTERNOON, 16 MAY 2005

PLAZA A, 1:20 TO 2:40 P.M.

Session 1pAOa

Acoustical Oceanography and Underwater Acoustics: Riverine Acoustics II

Robert Kieser, Cochair

Dept. of Fisheries and Oceans, Pacific Biological Station, 3190 Hammond Bay Rd., Nanaimo, BC V9T 6N7, Canada

Kenneth G. Foote, Cochair

Dept. of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, MA 02543

Invited Paper

1:20

1pAOa1. New single echo detection methods for shallow water fishery acoustics. Helge Balk, Torfinn Lindem, and Jan Kubecka (Dept. of Phys., PB1048 Blindern 0316 Oslo, Norway, helge.balk@fys.uio.no)

Acoustic target detection is commonly carried out with parametric single echo detectors. These detectors test one ping at a time and look for echoes fulfilling a set of criteria such as echo duration and shape. In shallow water, noise phenomena can distort echoes from fish and false fish echoes can be generated. This causes the parametric detector to produce fractionated tracks from fish surrounded by numerous noise detections. Parametric detectors utilize only small portions of the information available in a split beam echogram. By including information from more than one ping and from the background reverberation, a more robust fish detector has been designed. This detector, called the Cross Filter Detector (CFD), has now been further improved by applying the variance in the angle measurements.

1:40

1pAOa2. Echo characteristics of two salmon species. Patrick A. Neilson (School of Aquatic and Fishery Sci., Univ. of Washington, Box 355020, Seattle, WA 98195-5020), John K. Horne (Univ. of Washington, Seattle, WA 98195-5020), and Debby L. Burwen (Alaska Dept. of Fish and Game, Anchorage, AK 99518-1599)

The Alaska Department of Fish and Game relies on split-beam hydroacoustic techniques to estimate Chinook salmon (*Oncorhynchus tshawytscha*) returns to the Kenai River. Chinook counts are periodically confounded by large numbers of smaller sockeye salmon (*O. nerka*). Echo target-strength has been used to distinguish fish length classes, but was too variable to separate Kenai River chinook and sockeye distributions. To evaluate the efficacy of alternate echo metrics, controlled acoustic measurements of tethered chinook and sockeye salmon were collected at 200 kHz. Echo returns were digitally sampled at 48 kHz. A suite of descriptive metrics were collected from a series of 1,000 echoes per fish. Measurements of echo width were least variable at the -3 dB power point. Initial results show echo elongation and ping-to-ping variability in echo envelope width were significantly greater for chinook than for sockeye salmon. Chinook were also observed to return multiple discrete peaks from a single broadcast echo. These characteristics were attributed to the physical width of chinook exceeding half of the broadcast echo pulse width at certain orientations. Echo phase variability, correlation coefficient and fractal dimension distributions did not demonstrate significant discriminatory power between the two species. [Work supported by ADF&G, ONR.]

1:55

1pAOa3. Alaskan river environmental acoustics. Peter H. Dahl (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105-6698), Carl Pfisterer (Alaska Dept. of Fish & Game, Fairbanks, AK), and Harold J. Geiger (Alaska Dept. of Fish & Game, Douglas, AK)

Sonars are used by the Alaska Department of Fish and Game (ADF&G) to obtain daily and hourly estimates of at least four species of migratory salmon during their seasonal migration which lasts from June to beginning of September. Suspended sediments associated with a river's sediment load is an important issue for ADF&G's sonar operations. Acoustically, the suspended sediments are a source of both volume reverberation and excess attenuation beyond that expected in fresh water. Each can impact daily protocols for fish enumeration via sonar. In this talk,

results from an environmental acoustic study conducted in the Kenai River (June 1999) using 420 kHz and 200 kHz side looking sonars, and in the Yukon River (July 2001) using a 120 kHz side looking sonar, are discussed. Estimates of the volume scattering coefficient and attenuation are related to total suspended sediments. The relative impact of bubble scattering and sediment scattering is also discussed.

2:10

1pAOa4. Echo integration of nonuniform fish densities. Robert Kieser (Dept. of Fisheries and Oceans, Pacific Biological Station, 3190 Hammond Bay Rd., Nanaimo, BC, Canada V9T 6N7) and John Hedgpeth (Tenera Environ., San Luis Obispo, CA 93401)

Echo integration is a well recognized method for measuring backscatter intensity and for estimating fish density. EI is appropriate for high and low fish densities as long as the target distribution is uniform across the beam. This is generally the case in mobile applications that use a downward looking transducer. However uniform fish distribution across the beam cannot be assumed in riverine applications that use a stationary side looking system. Observed distributions are often very nonuniform especially when migrating fish are surface or bottom oriented. EI may still be possible if the relative vertical fish distribution is known and reasonably constant over time. *A priori* estimates of the vertical fish density distribution could be from split-beam observations when densities are lower or from video observations. A model for the EI of nonuniform fish densities is developed, typical results are simulated to test the model and its application is discussed.

2:25

1pAOa5. Acoustic quantification of fish in the riverine environment: Challenges. Kenneth G. Foote (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Quantifying fish acoustically in rivers presents many challenges. Some are common to other aquatic environments, but are exacerbated in rivers. Acoustic issues of particular concern are reviewed. These include the backscattering cross section of fish, sampling volume, and both volumetric and surface reverberation. Advantages of methods based on multiple observations of the same fish, such as Doppler analysis, target-tracking, and sensing the angular dependence of fish scattering, or on correlation analysis, given sufficient bandwidth, are emphasized. The need for calibration of involved acoustic devices is mentioned.

Session 1pAOB

Acoustical Oceanography: General Topics in Acoustical Oceanography

Michael A. Wolfson, Chair

*Applied Physics Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698**Contributed Papers*

3:00

1pAOB1. Measurements of the sound created by single bubbles fragmenting in turbulence. Grant Deane (Mail Code 0238, Scripps Inst. of Oceanogr., UCSD, La Jolla, CA 92093-0238)

One of the long-standing problems in wave noise is to understand the link between small-scale physical processes occurring in breaking wave crests and radiated noise. It has been known since the late 1980s that a broad-band component of wind-related noise comes from individual and collective bubble oscillations, but little progress has been made in quantitative calculations of noise levels from whitecaps using first-principles physics. Recent theoretical and laboratory studies have demonstrated that at least part of the wave noise spectrum (from a few hundred hertz up to a few kilohertz) is generated by bubbles fragmented by the turbulent fluid flow in the wave crest. Results from some laboratory studies of the sound radiated by isolated bubbles fragmenting in a turbulent jet as a simplified model for fragmentation in whitecaps will be presented and discussed. [Work supported by ONR.]

3:15

1pAOB2. Could animals detect the approaching tsunami? Srinivasan Jagannathan, Nicholas Makris (MIT, 77 Massachusetts Ave, 5-435, Cambridge, MA 02139, jsrini@mit.edu), and Purnima Ratilal (Northeastern Univ., Boston, MA 02115)

A number of sources report the phenomenon of animals running to safety well before the recent southeast Asian tsunami struck. The phenomenon was apparently observed on both sides of the Bay of Bengal, from Thailand to Sri Lanka. The latter is far from the epicenter of the earthquake. Did the earthquake supply the only warning or is it possible that the tsunami radiated seismo-acoustic signals of sufficient amplitude and frequency for animals to detect and interpret it long before it struck shore? This possibility is quantitatively investigated by solving the coupled hydrodynamic-acoustic equations with a tsunami driving term in a range-dependent ocean acoustic waveguide. Cases where the tsunami is in deep water and on the continental shelf are considered. [The general topic of this research follows from a suggestion made by Herman Medwin.]

3:30

1pAOB3. Estimation of bubble density with subharmonic acoustic wave in bubbly water. Byoung-Nam Kim, Kang Il Lee, Suk Wang Yoon (Dept. of Phys., SungKyunKwan Univ., Suwon 440-746, Republic of Korea, swyoon@skku.ac.kr), and Bok Kyoung Choi (Korea Ocean Res. and Development Inst., Republic of Korea)

Bubble density was estimated with a subharmonic acoustic wave generated in bubbly water. The subharmonic acoustic wave can be easily generated due to the nonlinearity of bubbly water if the frequency of primary acoustic wave is double of the bubble resonance frequency and the driving acoustic pressure amplitude exceeds a certain threshold value. The frequency of primary acoustic wave was varied from 200 kHz to 500 kHz while the bubble resonance radius at subharmonic frequency was from 12 μm to 28 μm . The pressure level of the subharmonic acoustic wave linearly increased as the driving acoustic pressure amplitude increased. With the subharmonic pressure level, the bubble density was estimated from nonlinear bubble oscillation equation [Yu. A. Ilinskii and E.

A. Zabolotskaya, J. Acoust. Soc. Am. **92**, 2837–2841 (1992)]. The estimated bubble densities were also compared with those from a linear conventional acoustic bubble sizing method. Bubble sizing with subharmonic acoustic wave seems to be easily utilized for the diagnosis of sandy sediment with bubbles.

3:45

1pAOB4. Acoustical estimation of scatterer population from total scattering measurements. Stephane G. Conti (Southwest Fisheries Sci. Ctr., 8604 La Jolla Shores Dr., La Jolla, CA 92037), Philippe Roux (Marine Physical Lab., La Jolla, CA 92093-0205), and David A. Demer (Southwest Fisheries Sci. Ctr., La Jolla, CA 92037)

Multifrequency backscattering measurements are more and more frequently used to estimate the populations of scatterers from echosounders data, and to separate the echoes from different species. Such analysis is based on the projection of the backscattered spectral data on the theoretical scattering spectra of the scatterers. Beside the accuracy of the theoretical predictions, the results of the inversion are conditioned by different parameters such as the ambient noise, the number of frequencies available, and the differences between the spectra for each scatterer. Here, we propose to evaluate the possibility of estimating the populations of scatterers from wide bandwidth total scattering cross section spectra measured in a reverberant media. The influences of the different parameters on the evaluated populations are investigated using simulations and experiments. Simulations were run either for similar spheres of different size, or different spheres of the same size. The experiments were performed in a reverberant tank with scatterers of different shapes having the same average total scattering cross section over the bandwidth.

4:00

1pAOB5. Modeling coherent Doppler sonar in fisheries acoustics. Cristina D. S. Tollefsen and Len Zedel (Dept. of Phys. and Phys. Ocean., Memorial Univ. of Newfoundland, St. John's, NL, Canada A1B 3X7, cristina@physics.mun.ca)

A computer model was developed to simulate the operation of a 250-kHz coherent Doppler sonar used to detect fish movements. The backscattered signal was constructed by summing contributions from many point targets. That signal was then detected and analyzed using the same method as used in the actual Doppler system. The model results reproduce predictions for the standard deviation of Doppler velocity estimates from volume backscatter based on the standard theoretical model of coherent pulse processing [Zrnic, IEEE Trans. Aerosp. Electron. Syst. **AES-13**, 344–354 (1977)]. However when the signal is modified to simulate the backscatter from a swimming fish (a single strong target among many weaker targets) the modeled standard deviation is 2 to 4 times lower than predicted by theory. Furthermore, the unusually low modeled standard deviations agree with laboratory and field observations. The model results confirm that the theoretical treatment used to predict the performance of Doppler measurements of current flow, blood flow, and atmospheric phenomena does not apply to the backscatter resulting from a single strong target such as a fish.

1pAOB6. Fish population dynamics revealed by instantaneous continental-shelf scale acoustic imaging. Purnima Ratilal (Northeastern Univ., 409 Dana Res. Ctr., Boston, MA 02115), Deanelle Symonds, Nicholas C. Makris (MIT, Cambridge, MA 02139), and Redwood Nero (Stennis Space Center, MS 39529)

Video images of fish population densities over vast areas of the New Jersey continental shelf have been produced from acoustic data collected on a long range bistatic sonar system during the Acoustic Clutter 2003 experiment. Areal fish population densities were obtained after correcting the acoustic data for two-way transmission loss modeled using the range-dependent parabolic equation, spatially varying beam pattern of the array, source level and mean target strength per fish. The wide-area fish density images reveal the temporal evolution of fish school distributions, their migration, as well as shoal formation and fragmentation at 50 s interval. Time series of the fish population within various density thresholds were made over the period of a day in an area containing millions of fish that at some instances formed a massive shoal extending over 12 km. The analysis shows that fish population in the area can be decomposed into a stable ambient population from lower-fish-density regions and a time-varying population composed from higher-density regions. Estimates of the differential speed between population centers of various shoals show that the average speed is on the order of a slow-moving surface vessel or submarine.

1pAOB7. Estimating biomass of densely populated fish schools from low-frequency acoustic scattering. Thomas R. Hahn (Rosenstiel School of Marine and Atmospheric Sci./AMP, Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149)

Fish schools of sufficient size and density can be efficient scatterers of low-frequency sound. This results from the large response of the fishes' swim bladders to incoming acoustic waves at frequencies at or close to their natural resonance frequencies. Acoustically, these fish schools can be viewed as a single object with acoustic bulk parameters determined from the properties of the individual fish. This effective medium approach incorporates the acoustical coupling of all fish due to multiple scattering and works for a wide range of fish-sizes, -numbers and -densities. If the contrast between the effective medium and the surrounding water is large enough, resonant enhancement occurs. These modes of collective oscillations have been conclusively observed by many authors in the context of bubble clouds. A study of the analytical structure of the effective-medium scattering amplitude reveals simple expressions relating the center frequencies and widths of low-frequency collective resonances to the overall size and fish number density of the fish school. If observed, these resonances might yield stable information that together with the mean resonance frequency of individual fish could lead to simple and practicable estimates of the total school biomass.

MONDAY AFTERNOON, 16 MAY 2005

REGENCY A, 1:30 TO 4:30 P.M.

Session 1pBB

Biomedical Ultrasound/Bioresponse to Vibration and Physical Acoustics: Tissue Response to Shock Waves II

Reiner Schultheiss, Cochair

SWS Shock Wave Systems AG, Wilen 4, 8574 Illighausen, Switzerland

Wolfgang Schaden, Cochair

Landstrasser Hauptstrasse 83, Vienna 1030, Austria

Invited Papers

1:30

1pBB1. Acoustic regulation of extracorporeal shock wave (ESW) therapy devices in the U.S. Subha Maruvada and Gerald R. Harris (Food and Drug Administration, Ctr. for Devices and Radiological Health, 12725 Twinbrook Pkwy., Rockville, MD 20852, subha.maruvada@fda.hhs.gov)

The focused, large amplitude pressure fields produced by ESW lithotripsy devices were shown in the early 1980s to provide an efficient means for fragmenting urinary tract calculi. More recently, orthopedic applications of intense pressure pulses for pain relief and fracture healing have been developed. Under the US Medical Device Amendments of 1976, ESW therapy devices were deemed Class III, meaning that a pre-market application typically would be supported by both pre-clinical and clinical studies. This classification still applies, except for ESW lithotripters indicated for fragmenting kidney and ureteral calculi. These devices were reclassified to Class II in 2000, resulting in a simpler path to market in which a demonstration of substantial equivalence to a currently marketed device is sufficient. As part of its regulatory responsibility to address the safety and effectiveness of these devices, particularly with regard to acoustic output, the US Food and Drug Administration has recognized two International Electrotechnical Commission (IEC) standards for ESW lithotripters, one covering field measurements (IEC 61846) and the other dealing with labeling and other safety aspects (IEC 60601-2-36). Although these standards were designed primarily for lithotripsy, the FDA has used them where applicable in the regulatory analysis of other ESW therapy devices.

1:45

1pBB2. The acoustic fields of shock wave therapy devices. Robin O. Cleveland and Parag V. Chitnis (Dept of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215, robinc@bu.edu)

We report measurements of a number of different shock wave therapy (SWT) devices. Two devices were electrohydraulic (EH): one had a large shock source (HMT Ossatron) and the other was a small hand-held source (HMT Evotron). The other device was a pneumatically driven device (EMS Swiss Dolorclast) and two different hand pieces were measured, one with an "unfocused" head and the other with a "focused" head. We found that the EH sources generated focused shock waves with a positive phase about 1 microsec long and peak pressure around 40 MPa, however, the acoustic output of the HMT Evotron appeared to be independent of the power setting of the machine. For the pneumatic source the duration of the positive phase was greater than 4 microsec and the peak pressure about 7 MPa. There was no clear shock front present and the waveform had a complex tail structure that was dependent on the power setting of the machine. We found that the focused hand-piece did not generate a focused acoustic field. The results are compared to reports of measurements from electromagnetic SWT devices. We contrast measurements made with different hydrophone systems: fiber-optic probe hydrophone, PVDF membrane hydrophone and PVDF bullet-shaped hydrophone.

2:00

1pBB3. Different shock front characteristics for non-urological treatments. Reiner Schultheiss (SWS Shock Wave Systems AG, Wilen 4, 8574 Illighausen, Switzerland)

Focused shock waves: Concrement disintegration and stimulation of Pseudarthroses gaps request the shock wave energy at limited spatial areas. High pressure amplitudes and energies need to exceed certain thresholds of stone material for mechanical disintegration work. Accordingly the generation of fractures within the medullar bone explained the stimulation of healing proportional to the acoustic energy. Recent clinical findings reveal identical or even better outcomes at low number of shocks. Biological model: Accepting the biological model as promoted by W. Schaden, the peak pressure and the energy density of the shock waves might be lowered dramatically. Activation of the body's healing mechanisms will be seen by ingrowth of new blood vessels and the release of growth factors. Unfocused shock wave sources: The biological model motivated the design of sources with low pressure amplitudes and energy densities. First: spherical waves generated between two tips of an electrode; and second: nearly even waves generated by generalized parabolic reflectors. Third: divergent shock front characteristics are generated by an ellipsoid behind $F2$. Unfocused sources are preferably designed for extended two dimensional areas/volumes like skin. Detailed acoustical parameters might be presented.

2:15

1pBB4. Tissue response to modified shock wave profiles. Pei Zhong, Yufeng Zhou, Franklin H. Cocks (Dept. of Mech. Eng. and Mat. Sci., Duke Univ., Box 90300, Durham, NC 27708), Charles Marguet, Glenn M. Preminger (Duke Univ.), and John Madden (Duke Univ.)

We have upgraded the original Dornier HM-3 lithotripter with a reflector insert to modify the lithotripter pulse profile and a piezoelectric annular array (PEAA) generator to produce microsecond tandem pulse lithotripsy. In this talk, we will present results from in vivo animal experiments using a swine model with or without surgically implanted artificial stones in the renal pelvis. Using this animal model, we have compared in vivo stone comminution and tissue injury, respectively, produced by the original versus the upgraded HM-3 under clinically relevant output settings. It was found that the upgraded HM-3 could produce significantly improved stone comminution with substantially reduced tissue injury, compared to the original HM-3. The underlying mechanisms that may contribute to the improved performance and safety of the upgraded HM-3 will be discussed. [Work supported by NIH DK52985 and DK58266.]

2:30

1pBB5. Reorganization of pathological control functions of memory—A neural model for tissue healing by shock waves. Othmar Wess (Storz Medical AG, Unterseestrasse 47, CH 8280 Kreuzlingen, Switzerland)

Since 1980 shock waves have proven effective in the field of extracorporeal lithotripsy. More than 10 years ago shock waves were successfully applied for various indications such as chronic pain, non-unions and, recently, for angina pectoris. These fields do not profit from the disintegration power but from stimulating and healing effects of shock waves. Increased metabolism and neo-vascularization are reported after shock wave application. According to C. J. Wang, a biological cascade is initiated, starting with a stimulating effect of physical energy resulting in increased circulation and metabolism. Pathological memory of neural control patterns is considered the reason for different pathologies characterized by insufficient metabolism. This paper presents a neural model for reorganization of pathological reflex patterns. The model acts on associative memory functions of the brain based on modification of synaptic junctions. Accordingly, pathological memory effects of the autonomous nervous system are reorganized by repeated application of shock waves followed by development of normal reflex patterns. Physiologic control of muscle and vascular tone is followed by increased metabolism and tissue repair. The memory model may explain hyper-stimulation effects in pain therapy.

2:45

1pBB6. The mechanism of shock wave treatment in bone healing. Ching-Jen Wang (Chang Gung Memorial Hospital at Kaohsiung, Taiwan, 123, Ta Pei Rd., Niao Sung Hsiang, Kaohsiung 833, Taiwan)

The purpose of this study was to investigate the biological mechanism of shock wave treatment in bone healing in rabbits. A closed fracture of the right femur was created with a three-point bend method and the fracture was stabilized with an intra-medullary pin. Shock waves were applied one week after the fracture. Twenty-four New Zealand white rabbits were randomly divided into 3 groups.

1p MON. PM

Group 1 (the control) received no shock waves; group 2 received low-energy; and group 3 high-energy shock waves. The animals were sacrificed at 24 weeks, and a 5-cm segment of the femur bone including the callus was harvested. The specimens were studied with histomorphological examination, biomechanical analysis and immunohistochemical stains. The results showed that high-energy shock waves improved bone healing with significant increases in cortical bone formation and the number neovascularization in histomorphology, better bone strength and bone mass in biomechanics, and increased expressions of angiogenic growth markers including BMP-2, eNOS, VEGF and PCNA than the control and low-energy shock wave groups. The effect of shock wave treatment appears to be dose-dependent. In conclusion, high-energy shock waves promote bone healing associated with ingrowth of neovascularization and increased expressions of angiogenic growth factors.

3:00

1pBB7. Working mechanism of extracorporeal shockwave therapy in non-urological disciplines. Wolfgang Schaden (Landstrasser Hauptstrasse 83, Vienna 1030, Austria, ismst@aon.at)

For 32 years of extracorporeal shockwave lithotripsy (ESWL) only the mechanical strength of shockwaves were of clinical interest. For use in orthopaedics, the absence of dangerous long term effects (malignant degeneration, etc.) is the only important message. The mechanical model tries to explain the effect of shock waves by the provocation of microlesions in the tissue stimulating repairing processes. First doubts on this mechanical model came up when Schaden (2001) could show, that less energy is more efficient in the treatment of non-unions. Due to the basic research of the last years knowledge increased about the microbiological effects. Under the influence of shock waves the change of permeability of cell membranes and the liberation of free radicals was reported. Also the production of nitric oxide (NO) and different growth factors like vascular endothelial growth factor (VEGF), bone morphogenetic proteins (BMP), transforming growth factor-beta 1 (TGF-b1), insulin-like growth factor-I (IGF-I) etc. was observed. The biological model tries to explain the effect of shock waves by stimulating the ingrowth of blood vessels and liberation of growth factors. Under the influence of shock waves, biological tissues seem to be able to produce important substances to initiate healing processes.

Contributed Papers

3:15

1pBB8. Effects of shock waves on growth of endothelial cells *in vitro*. Masaaki Tamagawa, Masanobu Kitayama, and Seiya Iwakura (Grad. School of Life Sci. and Systems Eng., Kyushu Inst. of Technol., Kitakyushu, Fukoka 808-0196, Japan, tama@life.kyutech.ac.jp)

Recently shock wave phenomena in living tissues are being widely applied in the fields of medical and chemical engineering, such as extracorporeal shock wave lithotripsy, bioprocess for environmental protection and tissue engineering. In the field of tissue engineering, the bone therapy to regenerate the bone by extracorporeal shock waves shows the possibility for new therapy. In this paper, to investigate the effects of shock waves on the endothelial cells *in vitro*, the cells by plane shock waves are observed by microscope and the growth rate and others are measured by image processing. The peak pressure works on the endothelial cells in water at the test case is 0.4 MPa. After working shock waves on suspended cells and fixed cells, the disintegration, shape and growth are investigated. It is found that the younger generation cells have small differences of shape index, and the growth rate of the shock-worked cells from 0 to 4 h are clearly high compared with control ones. It is concluded that once shock waves worked, some of them are disintegrated, but the other has capacity to increase growth rate of cell culture *in vitro*.

3:30

1pBB9. Study of a tissue protecting system for clinical applications of underwater shock wave. S. H. R. Hosseini and Kazuyoshi Takayama (Nanomedicine Div., Tohoku Univ., Biomed. Eng. Res. Organization, 2-1-1 Katahira, Aoba, Sendai 980-8577, Japan)

Applications of underwater shock waves have been extended to various clinical therapies during the past two decades. Besides the successful contribution of extracorporeal shock waves, tissue damage especially to the vasculature has been reported. These side effects are believed to be due to the shock wave-tissue interaction and cavitation. In the present research in order to minimize shock wave induced damage a shock wave attenuating system was designed and studied. The attenuating system consisted of thin gas packed layers immersed in water, which could attenuate more than 90% of shock waves overpressure. Silver azide micro-pellets (10 mg) were ignited by irradiation of a pulsed Nd:YAG laser to generate shock

waves. Pressure histories were measured with fiber optic probe and PVDF needle hydrophones. The strength of incident shock waves was changed by adjusting the distance between the pellets and the layers. The whole sequences of the shock wave attenuation due to the interaction of shock waves with the dissipating layers were quantitatively visualized by double exposure holographic interferometry and time resolved high speed photography. The attenuated shock had overpressure less than threshold damage of brain tissue evaluated from histological examination of the rat brain treated by shock waves.

3:45

1pBB10. Investigations on stone fragmentation in different extracorporeal shock wave lithotripsy sound fields *in vitro*. Thomas Dreyer, Marko Liebler, and Rainer Riedlinger (Universitaet Karlsruhe, IHE-Akustik, Kaiserstr. 12, D-76131 Karlsruhe, Germany, Thomas.Dreyer@ihe.uka.de)

The mechanism of stone fragmentation in ESWL applications is still under investigation. Devices showing a wide focal area and comparably low focal pressure amplitudes have been reported to disintegrate stones more efficiently as current clinical devices with high amplitudes and small focal areas. From this the question is raised whether the underlying different physical mechanisms or treatment issues, like stone localization and movement, are responsible for these results. In this paper fragmentation experiments *in vitro* with different stone types (e.g., HMT and BegoStone, 15 mm diam.) under different sound fields are presented. A self focusing piezoelectric transducer with a small focal area and peak pressure amplitudes of up to 125 MPa is used. The number of pulses was counted until a complete fragmentation through a 2 mm wire mesh is reached. In order to simulate wide-focus low-pressure conditions, the stones were placed in the prefocal region. Fragmentation results are compared to the case of focal placement. Initial breakage occurs earlier in the prefocal region for the HMT stones, whereas complete fragmentation is reached significantly earlier in the focus for all stone types.

1pBB11. A model for damage of microheterogeneous kidney stones. Andrew J. Szeri, Tarek I. Zohdi (Dept. of Mech. Engr., UC Berkeley, Berkeley, CA 94720-1740, Andrew.Szeri@Berkeley.edu), and John R. Blake (U. Birmingham, UK)

In this paper, a theoretical framework is developed for the mechanics of kidney stones with an isotropic, random microstructure—such as those comprised of cystine or struvite. The approach is based on a micromechanical description of kidney stones comprised of crystals in a binding matrix. Stress concentration functions are developed to determine load sharing of the particle phase and the binding matrix phase. As an illustration of the theory, the fatigue of kidney stones subject to shock wave lithotripsy is considered. Stress concentration functions are used to construct fatigue life estimates for each phase, as a function of the volume fraction and of the mechanical properties of the constituents, as well as the loading from SWL. The failure of the binding matrix is determined explicitly in a model for the accumulation of distributed damage. Also considered is the amount of material damaged in a representative non-spherical collapse of a cavitation bubble near the stone surface. The theory can be used to assess the importance of microscale heterogeneity on the comminution of renal calculi and to estimate the number of cycles to failure in terms of measurable material properties.

1pBB12. A mechanistic analysis of stone comminution in lithotripsy. Adam D. Maxwell (Ctr. for Industrial and Medical Ultrasound, APL, Univ. of Washington, Seattle, WA, 98105), Oleg A. Sapozhnikov (Moscow State Univ., Moscow, Russia), Michael R. Bailey, Brian MacConaghy, and Lawrence A. Crum (Univ. of Washington, Seattle, WA 98105)

In vitro experiments and an elastic wave model were employed to isolate and assess the importance of individual mechanisms in stone comminution in lithotripsy. Cylindrical U-30 cement stones were treated in an HM-3-style research lithotripter. Baffles were used to block specific waves responsible for spallation, squeezing, or shear. Surface cracks were added to stones to simulate the effect of cavitation, then tested in water and glycerol (a cavitation suppressive medium). Each case was simulated using the elasticity equations for an isotropic medium. The calculated location of maximum stress compared well with the experimental observations of where cracks naturally formed. Shear waves from the shock wave in the fluid traveling along the stone surface (a kind of dynamic squeezing) led to the largest stresses in the cylindrical stones and the fewest SWs to fracture. Reflection of the longitudinal wave from the back of the stone—spallation—and bubble-jet impact on the proximal and distal faces of the stone produced lower stresses and required more SWs to break stones. Surface cracks accelerated fragmentation when created near the location where the maximum stress was predicted. [Work supported by NIH DK43881, NIH-Fogarty, NSBRI SMS00203, RFBR, and ONRIFO.]

MONDAY AFTERNOON, 16 MAY 2005

GEORGIA A, 1:30 TO 2:50 P.M.

Session 1pEAa

Engineering Acoustics: Acoustic Transducers and Systems

Kim C. Benjamin, Chair

Naval Undersea Warfare Center, Newport, RI 02841

Chair's Introduction—1:30

Contributed Papers

1:35

1pEAa1. Seismic landmine detection using microphones as near-ground sensors. Gregg D. Larson, James S. Martin (Georgia Inst. of Technol., School of Mech. Eng., Atlanta, GA 30332), and Waymond R. Scott, Jr. (Georgia Inst. of Technol., Atlanta, GA 30332)

Seismic landmine detection systems interrogate the near-surface layers of the ground by propagating Rayleigh surface waves through the region of interest and remotely detecting surface displacements. However, in handheld or robotic applications, commercial microphones could be utilized as near-ground sensors. The primary surface wave for landmine detection is the Rayleigh surface wave, which propagates at subsonic speeds in typical soils. Therefore, the acoustic wave generated in the air by the Rayleigh wave is evanescent. In general, the Rayleigh waves acoustic pressure can only be accurately measured well within a seismic wavelength of the surface. As ambient acoustic noise and reverberation tend to decrease the signal-to-noise ratio, several techniques have been investigated to improve the measurements. The signal-to-noise ratio of microphone measurements is significantly improved by decreasing the microphone's height above the soil surface or by improving the coupling of the microphone to the evanescent field with a horn. Planar near-field acoustic holography has also been used to back-propagate these signals and calculate surface displacements; this can potentially enhance the performance by preferentially amplifying the effects of the Rayleigh wave. Measurements with microphones have detected both anti-personnel and anti-tank landmines and compare well with radar sensor measurements.

1:50

1pEAa2. Design and analysis of a silicon condenser microphone. Jin H. Huang (Dept. of Mech. and Comput.-Aided Eng., Feng Chia Univ., Taichung 407, Taiwan) and Han Kao (Penn State Univ., State College, PA 16804)

The design and analysis of a capacitive microphone is presented in this paper. Capacitive microphone with considerably higher sensitivity and low power consumption offers the innovative design for sound pressure micro-sensors. Polysilicon with smooth surfaces and incredible low residue stress is used for diaphragm of the capacitive microphone. Two methods—equivalent circuit method and finite element method—have been applied in this research to achieve the highest sensitivity of capacitive microphone. Optimal diaphragm edge width, thickness, and air gap have been determined through analyzing and simulating sound pressure micro-sensors, understanding the variation of geometry dimensions parameters between diaphragm and air gap distance, and dissecting the impact among sensitivity, nature frequency and electric field. Consequently, the valuable design model is able to be provided for the best choice of sound pressure micro-sensors among different materials. In addition, the results can improve and control the performance and dimensions of this microsensor. Furthermore, the microphone is fabricated using a combination of surface and bulk micromachining techniques which has favorable integrated capability of CMOS (Complementary Metal-Oxide Semiconductor). These device and techniques are promising for the future production.

2:05

1pEAa3. Mass loading effect of shock accelerometers. Lixue Wu, George S. K. Wong, Peter Hanes, and Won-Suk Ohm (Inst. for Natl. Measurement Standards, Natl. Res. Council Canada, Bldg. M-36, 1200 Montreal Rd., Ottawa, ON, Canada K1A 0R6)

Mass loading affects the sensitivity of an accelerometer. The mass loading effect can be corrected using mass loading correction curves published by manufacturers. These curves, however, are only applicable to the sinusoidal acceleration below 500 m/s². The mass loading effect on the sensitivity of an Endevco 2270 accelerometer in shock calibration, from 500 m/s² and up, was investigated using a laser vibrometer. A new method for conversion of a velocity signal to an acceleration signal was developed. With this method, the sensitivities of the above accelerometer for different mass loads at different shock levels were measured. The mass loading effect in shock calibration for this accelerometer was then obtained. The limitations of the sensitivity measurements were also studied. The variance of the measured sensitivity was mainly due to the resolution limitation (8-bit) of the A/D converter in the digital oscilloscope. A correlation matching algorithm was then developed that utilizes the similarity between the measured acceleration signal and that converted from the velocity signal to further improve the resolution of the digital oscilloscope.

2:20

1pEAa4. Characterization of a small moving-magnet electrodynamic linear motor for use in a thermoacoustic refrigerator. Jin Liu and Steven L. Garrett (Grad. Program in Acoust., Penn State, State College, PA 16804)

The mechanical and electrodynamic parameters of a small, potentially inexpensive, moving-magnet electrodynamic linear motor are determined experimentally. Employing the formalism introduced by Wakeland [J. Acoust. Soc. Am. **107**, 827–832 (2000)], these parameters are used to predict the electromechanical efficiency of the motor. The transduction coefficient Bl was observed to be a function of position. As will be shown analytically, the variation in Bl with position has a reduced effect on the

drivers output power because Bl is largest around the equilibrium position where the piston velocity is also largest. By mechanical co-linear joining of the armatures of two such motors, an electrodynamic load (dynamometer) is created. Motor efficiency is then measured as a function of electrical energy dissipated in resistors placed in series with the motor which acts as the alternator in the pair. The measured efficiencies are shown to be in good agreement with the predictions if a position-averaged effective transduction coefficient is introduced. Based on these results, this linear motor is judged to be an attractive power source in small electrically driven thermoacoustic refrigerator applications. [Work supported by the Applied Research Laboratory Exploratory and Foundational Research Fund and Ben and Jerrys with motors donated by Bose Corporation.]

2:35

1pEAa5. Radiated fields of thin and flexible high frequency loudspeakers. David A. Hutchins, Andrew P. Medley, and Duncan R. Billson (School of Eng., Univ. of Warwick, Coventry CV4 7AL, UK, D.A.Hutchins@warwick.ac.uk)

Thin and flexible electrostatic loudspeakers have been constructed, which can be curved into different shapes. This allows some interesting geometries to be explored, at frequencies in the range 10 kHz to 100 kHz. The research has compared the radiated fields measured experimentally with those predicted using established theory. The theoretical approach uses a new geometrical formulation to more efficiently predict radiated field patterns, where the curved surface is considered as a series of elements at specific angles to the beam axis. These elements can be of increased width compared to previous methods, making the approach computationally efficient. It is also a more robust approach than the traditional impulse response methods, in situations where the source size is comparable to the radiated wavelengths in air. Predictions could be made for both continuous and transient excitation. A comparison between theory and experiment has demonstrated that interesting field patterns can be generated. For instance, focussed sound fields can be produced using concave shapes, with predictable focal positions which are not at the center of curvature. These and other geometries will be described.

MONDAY AFTERNOON, 16 MAY 2005

GEORGIA A, 3:10 TO 4:30 P.M.

Session 1pEAb

Engineering Acoustics: Ducts, Barriers and Transmission Loss Testing

Stephen C. Thompson, Chair

Knowles Electronics Inc., 1151 Maplewood Dr., Itasca, IL 60143

Chair's Introduction—3:10

Contributed Papers

3:15

1pEAb1. Application of barrier in industrial noise control. Jonathan Chui, Yong Ma, and Salem Hertil (ATCO Noise Management, 1243 McKnight Blvd. NE, Calgary, AB, Canada)

Noise barriers have been widely used in environmental noise control, such as traffic and railway noise. Actually they are also cost-effective mitigation measures in industrial noise control. In this paper, the applications of noise barrier in power plant are introduced. Types of barrier and

barrier materials are briefly summarized and compared. A case study of noise barrier implement in a 50 MW power plant is presented. The plant is a natural gas-fired simple-cycle peaking facility, and consists of two opposed gas combustion turbine directly connected through a coupling to a single generator. Some residences are located around the facility. A noise barrier wall was designed and installed surrounding the facility to control the noise impact of the plant on the residences. The acoustic modeling software Cadna/A was used to predict the noise insertion loss of the barrier. The prediction results were also compared with the site measurements.

1pEAb2. Developing a procedure for transmission loss testing of noise control blankets. Corjan Buma (Dept. of Mech. Eng., Univ. of Alberta, Edmonton, AB, Canada T6G 2G8, bumacj@superiway.net)

The ASTM E90 test method allows variations of the basic procedure to accommodate diverse products (e.g., build a filler wall to test windows or doors). This paper describes the procedure developed to obtain a sound transmission loss rating for a line of noise-control blankets. The subject blankets are normally attached to an engine housing or used as a pipe lagging. It was originally intended to use the E90 procedure to see the increase in STC rating for a metal substrate without and with the blanket(s). However, this did not provide meaningful data. The sheet-metal substrate and the filler-wall test-opening were modified. Useful 1/3-octave band TL-data were then generated. Review of information publicly available for competitive products disclosed a distinct lack of useful acoustical data. The experience obtained in the study described could serve to further augment the E90 (or other) test procedure.

1pEAb3. A selected history of duct mode synthesis. Joe Posey (NASA Langley Res. Ctr., Hampton, VA 23681)

Attempts to attenuate noise propagating down a duct, with or without flow, must recognize that wall boundary conditions permit only certain spatial wave patterns (modes) to exist. The aircraft noise research community began building mode synthesizers at least thirty years ago in order to study duct propagation and radiation phenomena in a laboratory environment. Lockheed Georgia built a spinning mode synthesizer (SMS) for NASA in the 1970s. NASA used its SMS in a circular flow-duct to validate predictions of mode propagation through constrictions, lined sections, and inlets of various shapes. In the mid-1970s, Penn State created a mode synthesizer as part of a demonstration of active noise control in a circular, no-flow duct. NASA sponsored a series of studies in the 1990s aimed at maturing active control technology for ducted fan noise. Each of these active control systems was essentially a mode synthesizer coupled with a control system to cancel fan-generated noise. NASA is currently building a new mode synthesizer in a rectangular duct to study the effects of high-speed flow and curvature of the duct axis on advanced noise suppression technology. The evolution of mode-generation schemes will be discussed.

1pEAb4. Theoretical and experimental study of sound propagation in ducts with spatially periodic area changes. J. Ryan Nesbitt, Pavel V. Danilov, and Donald B. Bliss (Mech. Eng. and Mater. Sci., Duke Univ., Durham, NC 27705, dbb@duke.edu)

Sound propagation of one-dimensional waves through a tube with spatially periodic area changes is studied theoretically and experimentally. The resulting wave behavior is similar to that observed for periodic structures, with Bloch waves, pass-bands and stop-bands. The feasibility of this configuration is studied as a method to reduce sound radiation from noise sources having harmonics that fall in the stop bands. The performance of two such periodic passageways in series, each having different properties, is also studied. This approach to noise reduction is shown to give substantial attenuation. The bandwidth of the stop bands makes the design robust so that careful tuning is not required. Performance for broadband sound reduction is also evaluated. The relationship between area change parameters and the location and width of stop-bands is discussed. Experimental measurements show reasonable agreement with theoretical predictions as long as the actual geometry is accurately modeled.

1pEAb5. Usage of acoustical filters in transmission line enclosures as a replacement of fiber absorber. Onur Ilkorur, Ismail Yuksek, and Emre Omurlu (Yildiz Tech. Univ., Barbaros Bulv. 34349, Yildiz, Istanbul, Turkey)

The need for small-size loudspeaker enclosures is gaining importance, as the dimensions of the modern houses are getting smaller. Transmission line loudspeaker enclosures need more space than most other loudspeaker enclosures because of their design basics. In many commercial transmission line enclosures, fiber material is used for damping of the standing waves inside the line, resulting in a large size enclosure with a loss in sensitivity. However, their low frequency response is regarded as superior to some other enclosure types. In this paper, the frequency response of a transmission line enclosure is examined by replacing the fiber absorber with axial acoustical filters. It has been shown that, axial acoustical filters can perform as fiber absorber, requiring less space and improving the sensitivity. The other advantage of axial acoustical filters over fiber absorber is the simplicity of their electro-acoustic model. Measurements are performed using a Clio Standard System for fiber and axial acoustical filters. Electro-acoustic model and the measurement results are compared. The advantages of using axial acoustical filters are represented.

MONDAY AFTERNOON, 16 MAY 2005

GEORGIA B, 1:00 TO 5:10 P.M.

Session 1pED

Education in Acoustics: Acoustics Education for the Future

Daniel R. Raichel, Chair
2727 Moore Ln., Golden, CO 80506

Chair's Introduction—1:00

Invited Papers

1pED1. Future directions in musical acoustics education. Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115, rossing@physics.niu.edu)

The desire to understand sound production by musical instruments, transmission of musical sound to the listener, and the perception of musical sound has been a strong motivator for students to learn acoustics, and will no doubt continue to be. We briefly describe past and present practices in teaching musical acoustics and attempt to predict how the subject may develop in the future.

1:25

1pED2. The future of acoustics distance education at Penn State. Karen P. Brooks, Victor W. Sparrow, and Anthony A. Atchley (Penn State Grad. Program in Acoust., 217 Appl. Sci. Bldg., University Park, PA 16802)

For nearly 20 years Penn State's Graduate Program in Acoustics has offered a graduate distance education program, established in response to Department of Defense needs. Using satellite technology, courses provided synchronous classes incorporating one-way video and two-way audio. Advancements in technology allowed more sophisticated delivery systems to be considered and courses to be offered to employees of industry. Current technology utilizes real time video-streaming and archived lectures to enable individuals anywhere to access course materials. The evolution of technology, expansion of the geographic market and changing needs of the student, among other issues, require a new paradigm. This paradigm must consider issues such as faculty acceptance and questions facing all institutions with regard to blurring the distinction between residence and distance education. Who will be the students? What will be the purpose of education? Will it be to provide professional and/or research degrees? How will the Acoustics Program ensure it remains attractive to all students, while working within the boundaries and constraints of a major research university? This is a look at current practice and issues with an emphasis on those relevant to constructing the Acoustics Programs distance education strategy for the future.

1:45

1pED3. Acoustics in mechanical engineering undergraduate core courses: Challenges and opportunities. M. G. Prasad (Dept. of Mech. Eng., Stevens Inst. of Technol., Hoboken, NJ 07030, mprasad@stevens.edu)

Generally in an undergraduate curriculum of mechanical engineering, acoustics is not included as a core course. The major core courses deal with mechanics, design, dynamics of machinery, etc. However, engineering aspects of acoustics or noise can be included through elective courses. Given the limited slots for elective courses in a curriculum, it is difficult to run elective courses in acoustics regularly with a required number of students. The challenge is to find innovative ways to include acoustics into core courses so that all students are exposed to the field and its applications. The design and analysis of machine elements such as cams, gears, etc. are always part of core courses. It is in these contexts that the acoustics through noise aspects including multimedia can be introduced. Acoustics as an effect due to vibration as cause can be included in vibration analysis. A core course on system modeling can include acoustics. The integration of acoustical topics not only strengthens the core courses but also prepares the graduating engineer to deal with real problems better. Thus, it is important for academic acousticians to bring acoustics into the core courses. This paper presents some efforts to include the acoustics material in some core courses.

2:05

1pED4. Acoustics for the musically-gifted at Berklee College of Music. Anthony K. Hoover (Cavanaugh Tocci Assoc., Inc., 327 F Boston Post Rd., Sudbury, MA 01776)

Berklee College of Music has offered an undergraduate course in applied acoustics for eighteen years, and a growing number of students have chosen a career in acoustics. This paper will summarize some of the approaches used to convey meaningful information and methods, while also encouraging interest in acoustics, to a creative and energetic student population that traditionally avoids math and science. This paper will review the textbook developed for this class, the Acoustical Society At Berklee, and the annual Berklee Teachers On Teaching.

2:25

1pED5. Medical ultrasound education for bioengineers. Shahram Vaezy (Bioengineering and Appl. Phys. Lab., Box 355640, Univ. of Washington, Seattle, WA 98195)

The widespread adoption of ultrasound technologies in medicine has necessitated the development of educational programs to address the growing demand for trained expertise in both academia and industry. The demand has been especially great in the field of therapeutic ultrasound that has experienced a significant level of research and development activities in the past decade. The applications cover a wide range including cancer treatment, hemorrhage control, cardiac ablation, gene therapy, and cosmetic surgery. A comprehensive educational program in ultrasound is well suited for bioengineering departments at colleges and universities. Our educational program for students in Bioengineering at the University of Washington includes a year-long coursework covering theory and practice of ultrasound, conducting research projects, attending and presenting at weekly seminars on literature survey, presentations at scientific meetings, and attending specialized workshops offered by various institutions for specific topics. An important aspect of this training is its multi-disciplinary approach, encompassing science, engineering, and medicine. The students are required to build teams with expertise in these disciplines. Our experience shows that these students are well prepared for careers in academia, conducting cutting edge research, as well as industry, being involved in the transformation of research end-products to commercially viable technology.

2:45

1pED6. The elements of a comprehensive education for future architectural acousticians. Lily M. Wang (Architectural Engr. Prog., Univ. of Nebraska-Lincoln, 200B PKI, Omaha, NE 68182-0681, lwang4@unl.edu)

Curricula for students who seek to become consultants of architectural acoustics or researchers in the field are few in the United States and in the world. This paper will present the author's opinions on the principal skills a student should obtain from a focused course of study in architectural acoustics. These include: (a) a solid command of math and wave theory, (b) fluency with digital signal processing techniques and sound measurement equipment, (c) expertise in using architectural acoustic software with an understanding

of its limitations, (d) knowledge of building mechanical systems, (e) an understanding of human psychoacoustics, and (f) an appreciation for the artistic aspects of the discipline. Additionally, writing and presentation skills should be emphasized and participation in professional societies encouraged. Armed with such abilities, future architectural acousticians will advance the field significantly.

3:05–3:15 Break

3:15

1pED7. Meeting the textbook needs of modern acoustic courses. Daniel R. Raichel (Eilar Assoc., Encinitas, CA; and CUNY Grad. Ctr., NYC; 2727 Moore Ln., Fort Collins, CO 80526, draichel@comcast.net)

A truly modern textbook constitutes an essential tool for both the instructor and students of acoustics or, for that matter, of almost any other technological subject. Thus, a newer edition of an acoustics text for upper science/engineering undergraduates and graduate students should not only thoroughly cover the fundamentals (including the derivations of the wave equation for both solids and fluids—a feature, unfortunately, still lacking in many fundamental texts) but also include a coverage of the latest applications such as medical and industrial uses of ultrasound, computer programs for mapping noise contours and solving architectural acoustics problems, introduction to nonlinear acoustics, acousto-optics, sonoluminescence, voice recognition devices, surround-sound systems and newer means of sound reproduction such as iPod, advances in sound measurement equipment and prosthetic hearing devices, and so forth. While the metric system prevails throughout the world, the use of the British system of units in the example problems in a text still may be essential to U.S. students who are likely to work with architects or deal with non-SI units in U.S. industry and the military.

3:35

1pED8. Acoustics in the elementary classroom. Uwe J. Hansen (Indiana State Univ., Terre Haute, IN 47809)

The need for increased science exposure at all educational levels continues to be acute. Science is almost universally perceived as difficult, and its ability to raise the quality of life in the presence of apparently insurmountable social problems is increasingly suspect. Over the past 15 years we have conducted teacher workshops, visited classrooms, have organized hands-on demonstration sessions, judged science fairs, and mentored high school students in research efforts, all in an attempt to raise the level of enthusiasm for science. A look ahead suggests that the need continues. Elementary school teachers all too often limit their own science skills to plants and animals, and thus physics concepts do not get the exposure needed to generate the necessary excitement for the physical sciences. Workshops for Elementary grade teachers will be described, which are aimed at preparing teachers to use music as a vehicle to introduce basic physics concepts in the upper elementary grades.

Contributed Papers

3:55

1pED9. The elastodynamic Poynting vector bridges the gap in student understanding of complex wave phenomena. Cleon E. Dean (Phys. Dept., P.O.B. 8031, Georgia Southern Univ., Statesboro, GA 30460-8031, cdean@GeorgiaSouthern.edu) and James P. Braselton (Georgia Southern Univ., Statesboro, GA 30460-8093)

The advent of powerful graphic desktop computers and software allows the modern physics or engineering student to be shown dynamic physical processes formerly considered too complicated to present at the introductory level. For some time now desktop computers have been capable of showing intricate patterns of time evolving physical systems. As a simple example, students can be shown the energy flux of surface guided waves in both time averaged and time animated forms. Both vacuum and fluid loaded examples in various geometries are considered as examples. In an introductory wave phenomena class that considers electromagnetic, acoustic, and elastodynamic waves, a consistent approach using the Poynting vector field serves to bridge the gap between understanding of wave behavior for the three different types.

4:10

1pED10. Field studies in architectural acoustics using Tablet PCs. Daniel Boye (Phys. Dept., Davidson College, Davidson, NC 28035-7133)

Core requirements for the sciences within the liberal arts curriculum challenge students to become directly involved in scientific study. These requirements seek to develop scientifically literate leaders and members of society. Formal laboratory periods are not usually associated with these courses. Thus, conceptual discovery and quantitative experimentation must take place outside of the classroom. Physics 115: Musical Technology at Davidson College is such a course and contains a section dealing with architectural acoustics. Field studies in the past have been an awk-

ward and cumbersome activity, especially for non-science majors. The emerging technology of Tablet PCs overcomes many of the problems of mobile data acquisition and analysis, and allows the students to determine the locations of the rooms to be studied. The impulse method for determining reverberation time is used and compared with calculations based on room size and absorption media. The use of Tablet PCs and the publicly available freeware Audacity in field studies investigating architectural acoustics will be discussed. [Work supported in part by the Associated Colleges of the South through their Technology Fellowship program.]

4:25

1pED11. Developing technologies for bioacoustic vocal profiling as a viable component of integrative medical diagnostics and treatment. Sharry K. Edwards (Sound Health Res. Inst., Inc., Albany, OH 45710 and Capital Univ. of Integrative Medicine, Washington, DC)

Over the past 20+ years the pioneering field of Human Bioacoustics, which includes voice spectral analysis, has begun to model the frequencies and architecture of human vocalizations to identify the innate mathematical templates found within the various system of the human body. Using the idea that the voice is a holographic representation of health and wellness, these non-invasive techniques are being advanced to the extent that a computerized Vocal Profile, using a system of Frequency Equivalents, can be used to accurately quantify, organize, interpret, define, and extrapolate biometric information from the human voice. This information, in turn, provides the opportunity to predict, direct, and maintain intrinsic form and function. This novel approach has provided an accumulation of significant data but until recently has been without an efficient biological framework of reference. The emerging Mathematical Model being assembled through Human Bioacoustic research likely has the potential to allow Vocal Pro-

filing to be used to predict and monitor health issues from the very first cries of a newborn through the frequency foundations of disease and aging.

4:40

1pED12. Teaching sonar methods and technologies using a practical, real-life research environment. Tom Fedenczuk and Patricia Fryer (SOEST/HIGP, Univ. of Hawaii, 1680 East-West Rd., Honolulu, HI 96822)

The results and current developments of an educational outreach website and CD-ROM which allows students and instructors to follow the day-by-day progress of an ocean research expedition aboard the R/V Thomas G. Thompson are presented. Sonar concepts, techniques, and technologies are explained through daily updates, science objectives, data results, and lesson outlines. As the greatest source of data collected on the expedition, sonar technologies play a large role in the outreach effort. Students are exposed to hullmounted bathymetry data collection instrumentation that includes the EM300 and Hydrosweep, multibeam sonar systems. The concept of deep towed, side scan-sonar is introduced through the DSL-120 system (WHOI). Students are also introduced to sonar technologies incorporated into remotely operated vehicles like Jason II/Madea. These include the Imagex 855 scanning sonar, the RDI bottom-

tracking Doppler velocity logger, as well as navigational transponders. Each instrument and educational concept is detailed using video, Quick Time Virtual Reality, html text, and/or images. Using a Marisat satellite link, students are able to ask the researchers questions about their work and findings. [Work supported by a supplement to NSF/OCE-0002584 project: Collaborative Research: Studies of Deep-sourced Mud Volcanism in the Mariana Forearc: A DSL120, Jason ROV, and Coring Program.]

4:55

1pED13. Integration discipline acoustical engineering for training specialists in acoustics and close areas. Nickolay Zagrai and Vladislav Zakharevitch (TSURE, 347928, Taganrog, GSP-17a, Nekrasovskiy, 44, Russia, znp@tsure.ru)

The author's discipline acoustical engineering was developed. It is suggested to students before studying the core disciplines of professional programs. The contents of discipline are general ideas, requirements and possible realization of engineering decisions using acoustical fields in different areas of design. The principles of achieving goals in the area of acoustical engineering were methodically systematized and common concepts of their realization were formulated.

MONDAY AFTERNOON, 16 MAY 2005

BALMORAL, 1:30 TO 5:10 P.M.

Session 1pMU

Musical Acoustics, Architectural Acoustics and Psychological and Physiological Acoustics: Low Frequency Content in Music

Jonas Braasch, Cochair

McGill Univ., Faculty of Music, 555 Sherbrooke St., West, Montreal, H3A 1E3, Canada

William L. Martens, Cochair

McGill Univ., Faculty of Music, 555 Sherbrooke St., West, Montreal, H3A 1E3, Canada

Invited Papers

1:30

1pMU1. The impact of variation in low-frequency interaural cross correlation on auditory spatial imagery in stereophonic loudspeaker reproduction. William Martens (Faculty of Music, McGill Univ., 555 Sherbrooke St. W., Montreal, QC, Canada H3A 1E3)

Several attributes of auditory spatial imagery associated with stereophonic sound reproduction are strongly modulated by variation in interaural cross correlation (IACC) within low frequency bands. Nonetheless, a standard practice in bass management for two-channel and multichannel loudspeaker reproduction is to mix low-frequency musical content to a single channel for reproduction via a single driver (e.g., a subwoofer). This paper reviews the results of psychoacoustic studies which support the conclusion that reproduction via multiple drivers of decorrelated low-frequency signals significantly affects such important spatial attributes as auditory source width (ASW), auditory source distance (ASD), and listener envelopment (LEV). A variety of methods have been employed in these tests, including forced choice discrimination and identification, and direct ratings of both global dissimilarity and distinct attributes. Contrary to assumptions that underlie industrial standards established in 1994 by ITU-R. Recommendation BS.775-1, these findings imply that substantial stereophonic spatial information exists within audio signals at frequencies below the 80 to 120 Hz range of prescribed subwoofer cutoff frequencies, and that loudspeaker reproduction of decorrelated signals at frequencies as low as 50 Hz can have an impact upon auditory spatial imagery. [Work supported by VRQ.]

1:55

1pMU2. Loudspeaker and listener positions for optimal low-frequency spatial reproduction in listening rooms. David Griesinger (Lexicon, 3 Oak Park Dr., Bedford, MA 01730-1441)

This paper will briefly describe the physical and physiological mechanisms that enable low-frequency externalization and spatial reproduction in listening rooms. These mechanisms depend on reproducing a time-varying interaural time difference at the listening position through interference between symmetric and asymmetric room modes. The effect works successfully when the symmetric and

asymmetric room modes are driven by independent portions of a multi-channel signal, typically the in-phase and the anti-phase component of a two-channel recording. The paper will describe how this overlap can be optimized by adjusting the loudspeaker and listener positions for a variety of playback rooms. The results show that certain common room shapes yield low spatiality regardless of loudspeaker and listener positions.

2:20

1pMU3. Subjective comparison of single channel versus two channel subwoofer reproduction. Todd S. Welti (Harman Intl. Industries Inc., 8500 Balboa Blvd., Northridge, CA 91329)

Bass management has many advantages for surround sound listening, however there is still a question regarding audibility of two channel versus single channel bass reproduction. Many previous investigations have lacked rigor or been preliminary in nature. The study presented here includes strict control of nuisance variables and significance testing of results. A test is described wherein trained listeners compared four different subwoofer configurations in controlled listening tests, using a very sensitive triangle test, in a typical listening room. Methods of controlling nuisance variables are discussed. These include double blind testing, equalizing all responses flat below 120 Hz, and allowing pre-training for the listening test. Critical selection of audio test loops using low frequency decorrelation is discussed. Results are presented.

2:45

1pMU4. Physiological and content considerations for a second low frequency channel for bass management, subwoofers, and low frequency enhancement (LFE). Robert E. (Robin) Miller III (Filmmaker Technol., 606 W. Broad St., Bethlehem, PA 18018)

Perception of very low frequencies (VLF) below 125 Hz reproduced by large woofers and subwoofers (SW), encompassing 3 octaves of the 10 regarded as audible, has physiological and content aspects. Large room acoustics and vibrato add VLF fluctuations, modulating audible carrier frequencies to >1 Hz. By convention, sounds below 90 Hz produce no interaural cues useful for spatial perception or localization, therefore bass management redirects the VLF range from main channels to a single (monaural) subwoofer channel, even if to more than one subwoofer. Yet subjects claim they hear a difference between a single subwoofer channel and two (stereo bass). If recordings contain spatial VLF content, is it possible physiologically to perceive interaural time/phase difference (ITD/IPD) between 16 and 125 Hz? To what extent does this perception have a lifelike quality; to what extent is it localization? If a first approximation of localization, would binaural SWs allow a higher crossover frequency (smaller satellite speakers)? Reported research supports the Jeffress model of ITD determination in brain structures, and extending the accepted lower frequency limit of IPD. Meanwhile, uncorrelated very low frequencies exist in all tested multi-channel music and movie content. The audibility, recording, and reproduction of uncorrelated VLF are explored in theory and experiments.

3:10–3:20 Break

3:20

1pMU5. The influence of low frequency energy on listener envelopment. Gilbert A. Soulodre (Adv. Audio Systems, Commun. Res. Ctr., Ottawa, Canada K2H 8S2, gilbert.soulodre@crc.ca)

The importance of low frequencies to the perception of spatial impression in concert halls and multi-channel surround systems is well established. The present paper reports the results of a series of subjective tests that demonstrate the influence of low frequency lateral energy on the perception of listener envelopment. The results show that the laterally-arriving low frequency energy contributes significantly to listener envelopment, and new objective measures of listener envelopment are derived that account for this effect. The effect of the spatial distribution of low frequency energy on the perception of Bass is also discussed.

3:45

1pMU6. Modeling auditory localization in the low-frequency range. Jonas Braasch, William L. Martens, and Wieslaw Woszczyk (CIRMMT, Sound Recording Area, Faculty of Music, McGill Univ., Montreal, QC, Canada H3A 1E3, jb@music.mcgill.ca)

Binaural recordings were made for subwoofer reproduction of octave-band noise bursts at 31.5-Hz, 63-Hz and 125-Hz center frequencies, and these low-frequency responses were analyzed using a binaural model simulating human perception. As expected, the interaural level differences remained nearly constant for different sound source positions within this low-frequency range. On the basis of interaural time differences, however, the model was able to predict the left/right position of the sound source on the interaural axis. In order to visualize the cross-correlation peak at low frequencies in the ITD range from -1.5 ms to $+1.5$ ms, the cross-correlation functions were decompressed by taking them to the power of 40. At these low-frequencies, the range of phase difference does not vary much with different sound positions although the ITDs are on the same order as for higher frequencies (≈ -1.0 ms to 1.0 ms), but the human ability to resolve very small phase differences already has been shown in previous investigations. The predictions of the model simulation were verified in a listening test. The repetition of the experiment in a second more reverberant space showed similar reductions in performance for both the human listeners and the model. [Work supported by VRQ.]

4:10

1pMU7. The effects of acoustical treatment on lateralization of low-frequency sources. Timothy J. Ryan, William L. Martens, and Wieslaw Woszczyk (CIRMMT, Sound Recording Area, Faculty of Music, McGill Univ., Montreal, Canada)

Recently, the standard of using a single low-frequency driver in stereophonic sound reproduction systems has come into question. Though it is accepted that lateral discrimination and localization of signals is possible well into the subwoofer frequency range, the use of multiple subwoofers in small reverberant rooms remains of questionable value. While inter-aural level differences (ILDs) are negligible at low frequencies, source lateralization is possible at low frequencies by virtue of inter-aural time differences (ITDs). But when such reproduction is attempted in small rooms, strong early reflections and resonances associated with room modes can cause erroneous ITD information to be detected by a listener, thereby compromising a listener's ability to accurately locate the source of a low-frequency sound. Acoustical treatment can be employed to reduce the level of early reflections and low-frequency ringing associated with sharp resonant modes in small rooms. Such acoustical treatment often results in more accurate reproduction of ITDs, which enables more accurate localization of sound sources in the horizontal plane. This study investigated changes in measured interaural phase differences after a treatment scheme using both diaphragmatic and Helmholtz-style absorbers. The results show the viability of using multiple low-frequency drivers given adequate acoustical treatment of the reproduction space.

4:25

1pMU8. A variable passive low-frequency absorber. Niels Werner Larsen, Eric R. Thompson, and Anders Christian Gade (Acoust. Tech., Oersted DTU, Tech. Univ. of Denmark, Bldg. 352, DK-2800 Kgs. Lyngby, Denmark, nielswerner@enghavepark.dk)

Multi-purpose concert halls face a dilemma. They can host classical music concerts, rock concerts and spoken word performances in a matter of a short period. These different performance types require significantly different acoustic conditions in order to provide the best sound quality to both the performers and the audience. A recommended reverberation time for classical music may be in the range of 1.5–2 s for empty halls, where rock music sounds best with a reverberation time around 0.8–1 s. Modern rhythmic music often contains high levels of sound energy in the low frequency bands but still requires a high definition for good sound quality. Ideally, the absorption of the hall should be adjustable in all frequency bands in order to provide good sound quality for all types of performances. The mid and high frequency absorption is easily regulated, but adjusting the low-frequency absorption has typically been too expensive or requires too much space to be practical for multi-purpose halls. Measurements were made on a variable low-frequency absorber to develop a practical solution to the dilemma. The paper will present the results of the measurements as well as a possible design.

4:40

1pMU9. Sensitivity to intermodal asynchrony between acoustic and structural vibrations. Kent Walker and William L. Martens (McGill Univ., Faculty of Music, Sound Recording, 555 Sherbrooke Ouest, Montreal, QC, Canada, H3A 1E3, kent.walker@mail.mcgill.ca)

The purpose of this study was to discover the attributes of musical stimuli which facilitate sensory integration in bi-modal music reproduction systems incorporating sound and whole-body vibration. It was hypothesized that subjective judgments regarding bimodal synchrony would vary depending on the spectral, temporal, and spatial properties of the stimuli. To test this hypothesis, musical instruments with significant low frequency energy and a variety of spectra-temporal envelopes were recorded. These stimuli were then reproduced with varying intermodal delay and overlap in frequency content between displayed vibratory and acoustic components. The air-born component of the bimodal stimuli was presented via a multichannel loudspeaker array, with a direct sound component, as well as a reproduced indirect sound arriving from all around the observer. Psychometric functions were constructed for time order judgment (TOJ) over a range of intermodal delay values. Changes in the slope and intercept of the transformed psychometric functions gave a clear picture of the influence of spectra-temporal and spatial parameters of the multimodal stimuli, the most striking results being the decreased tolerance for intermodal asynchrony associated with instruments recorded in reverberant environments. [Work supported by a Grant from VRQ of the Government of Quebec.]

4:55

1pMU10. Low-frequency interaural cross correlation discrimination in stereophonic reproduction of musical tones. Sungyoung Kim and William L. Martens (Faculty of Music, McGill Univ., 555 Sherbrooke St. West, Montreal, QC, Canada H3A 1E3, sungyoungk@hotmail.com)

By industry standard (ITU-R. Recommendation BS.775-1), multichannel stereophonic signals within the frequency range of up to 80 or 120 Hz may be mixed and delivered via a single driver (e.g., a subwoofer) without significant impairment of stereophonic sound quality. The assumption that stereophonic information within such low-frequency content is not significant was tested by measuring discrimination thresholds for changes in interaural cross-correlation (IACC) within spectral bands containing the lowest frequency components of low-pitch musical tones. Performances were recorded for three different musical instruments playing single notes ranging in fundamental frequency from 41 Hz to 110 Hz. The recordings, made using a multichannel microphone array composed of five DPA 4006 pressure microphones, were processed to produce a set of stimuli that varied in interaural cross-correlation (IACC) within a low-frequency band, but were otherwise identical in a higher-frequency band. This correlation processing was designed to have minimal effect upon other psychoacoustic variables such as loudness and timbre. The results show that changes in interaural cross correlation (IACC) within low-frequency bands of low-pitch musical tones are most easily discriminated when decorrelated signals are presented via subwoofers positioned at extreme lateral angles (far from the median plane). [Work supported by VRQ.]

Session 1pPA

Physical Acoustics and Signal Processing in Acoustics: Multiple Stochastic Scattering of Elastic and Seismic Waves

Richard L. Weaver, Chair

Dept. of Theoretical and Applied Mechanics, Univ. of Illinois, 104 South Wright St., Urbana, IL 61801

Chair's Introduction—1:00

Invited Papers

1:05

1pPA1. From seismology to oceanography: Locating the sources of the Earth's hum. Barbara Romanowicz and Junkee Rhie (Berkeley Seismological Lab., 215 McCone Hall, Berkeley, CA 94720)

The observation of continuously excited free oscillations of the Earth, in the absence of earthquakes, was first made by Japanese scientists in 1998. Since then, attention has focused on elucidating the physical mechanism responsible for them. The mechanism must be shallow, as fundamental modes appear to be preferentially excited and it shows seasonal variability. An array-based method has been developed to detect and locate sources of very long period surface wave energy, utilizing the dispersive properties of Rayleigh waves and data from two large aperture arrays of very long period seismometers, in California and in Japan. It is shown that, for each array, there is a well defined preferential direction, which is stable over one season but changes significantly from winter to summer. The fluctuations as a function of time of the maximum stack amplitudes are correlated across the two arrays and point to the northern Pacific Ocean in the northern hemisphere winter and the southern oceans in the summer, correlating with the distribution of maximum wave height. It is inferred that the background oscillations originate primarily in the oceans, and are caused by a non-linear coupling mechanism involving the atmosphere (winds), the oceans (infragravity waves) and the seafloor.

1:30

1pPA2. Correlation in ambient seismic noise and the reconstruction of Green function. Michel Campillo, Laurent Stehly (Observatoire de Grenoble, BP 53, 38041 Grenoble, France, Michel.Campillo@ujf-grenoble.fr), Nikolai Shapiro, and Mike Ritzwoller (Univ. of Colorado, Boulder, CO)

Cross-correlations between long continuous records of ambient seismic noise at distant stations are investigated. The dominant part of the Green function, namely Rayleigh waves, are reconstructed in a broad period range. This property reminds of the fluctuation-dissipation theorem that relates the random fluctuations of a linear system and the system's response to an external force. Ambient seismic noise is indeed not a thermal noise but it can be considered as a random and isotropic wave field both because the distribution of the ambient sources responsible for the noise randomizes when averaged over long periods and because of scattering from heterogeneities that occur within the Earth. The dispersion curves of Rayleigh waves for the paths between the stations are measured from the correlations. On paths where direct measurements between earthquake and station are available, we show that they are in good agreement with those deduced from noise correlation. The measurement of correlation along paths crossing different geological structures allows to differentiate them, opening the way for a passive imaging of the Earth structure. The dispersion measurements are applied to seismic tomography at the regional scale. They make it possible to image crustal structures with a resolution higher than conventional techniques.

1:55

1pPA3. Seismic interferometry, with applications in passive reflection imaging. Kees Wapenaar and Deyan Draganov (Dept. of Geotechnology, Delft Univ. of Technol., P.O. Box 5028, 2600 GA Delft, The Netherlands, c.p.a.wapenaar@citg.tudelft.nl)

Seismic interferometry is the process of generating new seismic responses by crosscorrelating seismic observations at different receiver locations. A first version of this principle was derived in 1968 by Claerbout, who showed that the reflection response of a horizontally layered medium can be synthesized from the autocorrelation of its transmission response. Later he conjectured a similar principle for crosscorrelations of 3-D wave fields. In a similar fashion, Schuster (2001) introduced the principle of interferometric imaging, i.e., forming an image of the subsurface from crosscorrelated seismic traces. In this paper we first discuss the theory of seismic interferometry for arbitrary 3-D inhomogeneous media (deterministic or random). Starting with the Rayleigh-Betti reciprocity theorem and the principle of time-reversal, we derive a number of relations that form the basis for seismic interferometry (amongst others these relations prove Claerbout's conjecture). Despite the difference in assumptions, these relations show a close resemblance with those of Weaver and Lobkis (2001) for the retrieval of the Greens function from diffuse wave field correlations. Next we discuss a number of applications, like passive seismic reflection imaging, surface wave reconstruction, improving sparse data sets and interferometric imaging for different geometries.

1pPA4. Interferometric imaging and the principle of stationary phase. Roel Snieder (Ctr. for Wave Phenomena, Colorado School of Mines, Golden, Co 80401-1887)

Interferometric imaging where the Green's function is constructed using the correlation of complex wavefields recorded at two receivers is a rapidly emerging field. The methodology has been justified based on assumptions of equipartitioning of the normal modes of the system, as well as on various versions of the representation theorem applied to time-reversed waves. I will present another point of view that is based on stationary phase arguments. This complementary formulation of interferometric imaging gives insight in the physics of the emergence of the Green's function, and of the limitations that are encountered when applying this technique to data. Issues that will be covered are the relation between ensemble averaging and time averaging, and interferometric imaging of reflected waves. This example is of particular interest for exploration seismology since it shows that interferometric imaging may introduce spurious multiple reflections. Given the recent effort in "multiple suppression" in exploration seismology, the introduction of spurious multiple reflections is an undesirable artifact.

Contributed Paper

2:45

1pPA5. Diffuse fields in open systems and the emergence of the Green's function. Richard Weaver (Dept. of Theoretical and Appl. Mech., Univ. of Illinois, Urbana, IL 61801)

As is now well known, the relation between diffuse field correlations and the Green's function follows directly from a definition of a diffuse field as an uncorrelated smooth spectral superposition of normal modes. Such a definition is, however, inapplicable in most open structures, the earth in particular. A preferable definition might be that of room acoustics:

a diffuse field is an uncorrelated isotropic superposition of plane waves. But that definition is inapplicable to heterogeneous structures, or near boundaries. Here, a definition of a local diffuse field applicable to open heterogeneous systems is proposed. A local diffuse field is taken to be one in steady-state equilibrium with the field in a homogeneous region having an uncorrelated isotropic superposition of incident plane waves. This definition is applicable to both heterogeneous and open systems, and is shown using a reciprocity argument to lead to the familiar identity between the local Green's function of the structure and the diffuse fields correlations.

3:00–3:20 Break

Invited Papers

3:20

1pPA6. Phase statistics of multiply scattered ultrasonic waves in dynamic mesoscopic systems. John H. Page, Michael L. Cowan^{a)} (Dept. of Phys. and Astron., Univ. of Manitoba, Winnipeg, MB, Canada R3T 2N2), and Bart Van Tiggelen (CNRS/Universite Joseph Fourier, 38042 Grenoble, France)

In weakly scattering materials, detecting motion by measuring the change in phase of reflected ultrasonic waves forms the basis of the well-known technique of Doppler ultrasound. In strongly scattering media, these methods break down and the technique of diffusing acoustic wave spectroscopy (DAWS) was developed [Cowan *et al.*, Phys. Rev. Lett. **85**, 453 (2000)]. To explore the use of phase information to investigate the dynamics of multiply scattering media, the temporal fluctuations in the phase of ultrasonic waves transmitted through a time-varying mesoscopic sample have been measured. We have compared phase statistics and correlations to detailed theoretical predictions based on circular Gaussian ($C1$) statistics [Genack *et al.*, Phys. Rev. Lett. **82**, 412 (1999)]. So far, excellent agreement is found. The cumulative phase is found to undergo a Brownian type process, described by a phase diffusion coefficient. A fundamental relationship between the variance in the phase of the transmitted waves and the fluctuations in the phase of individual scattering paths is predicted theoretically and verified experimentally. This relationship not only gives deeper insight into the physics of the phase of multiply scattered waves, but also provides a new, mesoscopic way of probing the motion of the scatterers in the sample. ^{a)}Currently at Department of Physics, University of Toronto, Toronto, ON, Canada M5S 3E3

3:45

1pPA7. Evolution of diffuse fields in heterogeneous slabs. Joseph A. Turner and Goutam Ghoshal (Dept. of Eng. Mech., Univ. of Nebraska-Lincoln, Lincoln, NE 68588)

Fundamental studies of elastic wave scattering in heterogeneous media are applicable for problems at several length scales from ultrasonic to seismic waves. The intermediate scattering regime that lies between the single scattering and the diffusion limits is perhaps the least understood. In this presentation, both the steady-state and time dependent scattering problems are examined for this regime within the context of radiative transfer theory. The focus here is on slab geometries for which the scattering medium lies between two parallel boundaries separated by a distance investigated over a range from several to tens of mean free paths. The spatial distribution, temporal evolution, and partitioning of the diffuse longitudinal and shear energies are studied as a function of direction and frequency for several types of microstructure including polycrystalline metals, concrete, porous media, and geophysical media. The longitudinal and shear flux reflected from and transmitted through the slab are also discussed due to their importance for experimental materials characterization. Finally, comparisons are made with direct numerical scattering simulations for some of the microstructures considered. The results are anticipated to shed insight on this important intermediate scattering regime. [Work supported by NSF and DOE.]

1pPA8. The transition to equipartitioning and its relation to scattering strength. Alison E. Malcolm, John A. Scales (Colorado School of Mines, Golden, CO 80401, amalcolm@dix.mines.edu), and Bart A. van Tiggelen (CNRS/Université Joseph Fourier-Grenoble, France)

In a coarse-grained rock scattering plays a dominant role in determining the characteristics of the ultrasonic wavefield. Immediately after a pulse excites waves in such a sample energy propagates away from the source as a packet. As this packet travels, the individual grains scatter its energy in all directions. This scattering process continues until there are equal amounts of energy propagating in all directions. If the scattering is strong enough, this transition is completed before the signal strength has relaxed to the ambient level. By collecting a dense data set on several rock samples we are able to watch this transition and see the influence of grain size on the wavefield. More quantitatively we show, with noncontacting laboratory data, how this transition can be tracked using the symmetry of the advanced and retarded Green functions. These Green functions are estimated by cross-correlating the signal recorded at two detectors; the two Green functions have equal amplitude only when equal amounts of energy are propagating in all directions. [Work supported by the NSF (EAR-0111804), the US Army Research Office (DAAG55-98-1-0070), the Center for Wave Phenomena, and ACI 2066 of the French Department for Research.]

Contributed Paper

4:35

1pPA9. Convergence rates for diffuse field-field correlations. Richard Weaver (Dept. of Theoretical and Appl. Mech., Univ. of Illinois, Urbana, IL 61801)

A model of diffuse fields in open structures is taken to be that resulting from a Gaussian random distribution of sources spread over all space. It is shown, consistent with recent literature, that this ensemble of fields gives rise to a field-field correlation function R between two points that has

expectation $\langle R \rangle$ equal to the Green's function between those points. It is furthermore found that this model lends itself to calculations of the variance of R , and thus to estimates of the degree to which an R calculated using finite amounts of data will conform to the Green's function. The model indicates, in accord with observations, that such conformation is strongest at low frequencies. Ray arrivals are detectable only if sufficient data has been collected; the amount of data needed scales with the square of the source-receiver separation, and the square of the frequency. Applications to seismology are discussed.

MONDAY AFTERNOON, 16 MAY 2005

REGENCY C, 1:30 TO 4:50 P.M.

Session 1pPP

Psychological and Physiological Acoustics: Auditory Perception and Psychophysics I

Elizabeth A. Strickland, Chair

Purdue Univ., Speech Language and Hearing Science, 500 Oval Dr., West Lafayette, IN 47907-2038

Chair's Introduction—1:30

Contributed Papers

1:35

1pPP1. Fear conditioning facilitates rats gap detection measured by prepulse inhibition of the startle reflex. Dan Zou, Xihong Wu, and Liang Li (Dept. of Psych., Natl. Key Lab. on Machine Percept. Speech and Hearing Res. Ctr., Peking Univ., Beijing, 100871 China)

A low-intensity acoustic event presented shortly before an intense startling sound can inhibit the acoustic startle reflex. This phenomenon is called prepulse inhibition (PPI), and is widely used as a model of sensorimotor gating in both humans and animals. Particularly, it has been used for evaluating the aging effect on the mouse's ability to detect a silent gap in otherwise continuous sounds. The present study extended this model to the emotional modulation of gap detection. The results show that a silent gap embedded in each of the two broadband noise sounds (55 dB SPL), which were delivered by two spatially separated loudspeakers, could inhibit the startle reflex that was induced by a loud sound presented from the third loudspeaker 50 ms after the gap. The inhibitory effect largely depended on the duration of the gap, with the mean duration threshold around 11 ms across 18 rats tested. Pairing the gap with foot shock in a temporally specific manner, but not in a temporally random

manner, significantly reduced the duration threshold. Thus this study established a new animal behavioral model both for studying auditory temporal processing and for studying auditory signal-detection plasticity induced by emotional learning.

1:50

1pPP2. Hearing levels in US adults aged 20–69 Years: National Health and Nutrition Examination Survey 1999–2002. William J. Murphy, Christa L. Themann, and John R. Franks (Hearing Loss Prevention Team, NIOSH, MS C-27, 4676 Columbia Pkwy., Cincinnati, OH 45226-1998)

The National Health and Nutrition Examination Survey (NHANES) is a nationally representative, population-based survey designed to assess the health and nutritional status of the civilian, non-institutionalized US population. Data were collected through a personal interview regarding health history and through physical examination. Earlier NHANES surveys were conducted on a periodic basis; however, in 1999, NHANES began collecting data on a continuing, annual basis. During NHANES I, which ran from 1971–1975, audiometric testing was conducted on adults aged 25–74

years. No subsequent testing of adults was conducted in the NHANES program until 1999, when NHANES began audiometric testing of adults aged 20–69 years. This report examines the hearing levels for adults in the United States and compares them with the hearing data from NHANES I. Hearing levels are grouped by age and are grouped by ethnicity and gender.

2:05

1pPP3. Application of musical timbre discrimination features to active sonar classification. Victor W. Young, Paul C. Hines, and Sean Pecknold (Defence R&D Canada–Atlantic, P.O. Box 1012, Dartmouth, NS, Canada B2Y 3Z7, victor.young@drdc-rddc.gc.ca)

In musical acoustics significant effort has been devoted to uncovering the physical basis of timbre perception. Most investigations into timbre rely on multidimensional scaling (MDS), in which different musical sounds are arranged as points in multidimensional space. The Euclidean distance between points corresponds to the perceptual distance between sounds and the multidimensional axes are linked to measurable properties of the sounds. MDS has identified numerous temporal and spectral features believed to be important to timbre perception. There is reason to believe that some of these features may have wider application in the disparate field of underwater acoustics, since anecdotal evidence suggests active sonar returns from metallic objects sound different than natural clutter returns when auralized by human operators. This is particularly encouraging since attempts to develop robust automatic classifiers capable of target-clutter discrimination over a wide range of operational conditions have met with limited success. Spectral features relevant to target-clutter discrimination are believed to include click-pitch and envelope irregularity; relevant temporal features are believed to include duration, sub-band attack/decay time, and time separation pitch. Preliminary results from an investigation into the role of these timbre features in target-clutter discrimination will be presented. [Work supported by NSERC and GDC.]

2:20

1pPP4. Frequency-time dispersion products revisited. William S. Hellman (Dept. of Phys. and Hear. Res. Ctr., Boston Univ., Boston, MA 02115, hellman@buphy.bu.edu) and H. Steven Colburn (Hear. Res. Ctr. and Dept. Biomed. Eng., Boston, MA 02115)

The interaction of bandwidth and duration is an important aspect of signal choice in psychophysical experiments. The traditional measures yield frequency-time dispersion products which have a lower bound of 0.25. Gabor [J. Inst. Elect. Eng. Part III **93**, 429–457 (1946)] argued that the measures for the standard deviation and mean in the frequency domain do not yield formulas in agreement with intuition when applied to the calculation of frequency dispersions for real signals $f(t)$. Gabor's prescription for correcting the problem was to limit the integrations in frequency statistic calculations to the positive frequency domain and, for consistency, use the analytic signal corresponding to $f(t)$ to calculate time dispersions. A more recent statement of this prescription can be found in L. Cohen [*Time-Frequency Analysis*, (Prentice Hall, 1995)]. However the elimination of the negative frequencies introduces difficulties which have not been fully addressed in the literature. It is shown here, that unless the Fourier transform of the signal vanishes at zero frequency, a linear divergence will appear in the time dispersion derived from the analytic signal. In such instances, $f(t)$ must be used to calculate the time dispersion. However the restriction to positive frequencies can generate frequency-time dispersion products which fall below the expected lower limit of 0.25.

2:35

1pPP5. Diotic and dichotic discrimination of binary sequences. Stanley Sheft, William A. Yost, and Raymond H. Dye (Parmlly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, ssheft@luc.edu)

Binary-sequence discrimination was compared for diotic and dichotic stimuli. Sequences consisted of 4 to 32 wideband-noise pulses with pulse duration ranging from 8 to 32 ms. Diotic sequences were distinguished by pulse-amplitude pattern, while dichotic patterns differed by their sequence

of ear of presentation. Discrimination was measured as a function of the number of pattern elements that differed between the standard and comparison sequences with temporal location of the altered pulses randomly selected on each trial. Additional fringe pulses bracketed the target sequences to avoid onset and offset cuing. Neither diotic nor dichotic performance was monotonic with the ratio of the number of altered to sequence pulses, with greater exception noted in the dichotic results. Except at the shortest pulse duration, diotic performance was significantly better than that obtained in the dichotic condition with similar pulse duration and numbers of altered and sequence pulses. For the range of stimulus parameters used, sequence discrimination often relied on a global percept rather than processing of individual pulse attributes with timbre differences cuing diotic discrimination. Though exhibiting fine resolution, results suggest poorer ability of the binaural than monaural system at extracting a global percept to cue sequence discrimination. [Work supported by NIDCD.]

2:50

1pPP6. Spatial release from masking for amplitude modulated and non-modulated noise stimuli. Norbert Kopčo (Hearing Res. Ctr., Boston Univ. and Tech. Univ., Košice, Slovakia) and Barbara G. Shinn-Cunningham (Boston Univ.)

The ability to hear a target sound (T) masked by another sound (M) improves when the T and M are spatially separated, a phenomenon known as spatial release from masking (SRM). Target detectability is also influenced by temporal characteristics of T and M (e.g., by the presence or absence of amplitude modulation). The current study examines how SRM is influenced by amplitude modulation. Detection thresholds were measured for a broadband noise target (T) temporally and spectrally centered within a broadband noise masker (M). Thresholds were measured for all combinations of five spatial configurations of T and M and five modulation conditions (all combinations of T and M modulated and unmodulated; when both T and M were modulated, the modulation could either be equal or pi out of phase). In all cases, the amplitude modulation, if present, had a rate of 40 Hz and depth of 0.5. Modulation had a complex effect on detection threshold. Thresholds improved by as much as 6 dB (relative to the no-modulation control) in some spatial configurations, but were nearly unaffected in others. These results have important implications for understanding the processes involved in the perception of simultaneous complex signals. [Work supported by NSF and NAS.]

3:05–3:20 Break

3:20

1pPP7. Loudness estimation in the presence of vertical vibrations. Etienne Parizet and Benjamin Marpe (LVA–Insa Lyon 25 bis, avenue Jean Capelle F-69621 Villeurbanne Cedex, France)

The goal of the study was to check whether vibrations submitted to a subject can modify his loudness evaluation. For that purpose, the subject was seated on a rigid chair vertically moved by a shaker, at the frequency of 28 Hz and at 6 different levels. He was also exposed to a pure sound: either a 28 Hz tone produced by a subwoofer, or a 1000 Hz tone produced by headphones. For each combination of sound and vibration levels, the subject had to estimate the magnitude of the loudness of the tone, as compared to a reference tone, heard without any vibration excitation. He also had to evaluate the magnitude of vibration using a 10 point scale. Twenty subjects participated in the experiment. The results showed that vibrations did not influence loudness estimation. On the other hand, the estimation of vibration level was significantly influenced by the level of the 28 Hz pure tone, which can indicate an interaction between the perception of sound by the body and the perception of vibrations.

3:35

1pPP8. Temporal and spectral interaction in loudness perception. Benjamin Pedersen and Wolfgang Ellermeier (Sound Quality Res. Unit, Dept. of Acoust., Aalborg Univ., Fredrik Bajers Vej 7 B5, DK-9220 Aalborg, Denmark)

An experiment was conducted to investigate how changes in spectral content influence loudness judgments. Six listeners were asked to discriminate sounds, which were of one second duration and changing in level every 0.1 s. In one condition the first half of the sound was low-pass filtered and the second half high-pass filtered. In a second condition the opposite order was used. In a third condition no filtering was applied and the frequency spectrum was simply white noise. The results were analyzed using a statistical method, which assigns relative weights to the ten temporal segments. In this way individual weighting curves were obtained for each condition. Listeners tended to emphasize the beginning of the sound in their loudness judgments. When the frequency spectrum changed in the middle of the sound, however, the weighting of the onset of the new spectral content was emphasized as well. This outcome is inconsistent with overall temporal integration, and argues for a cognitive mechanism allocating attention to changes in an event sequence.

3:50

1pPP9. Spectral resolvability of iterated rippled noise. William A. Yost (Parmly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626)

A forward-masking experiment was used to estimate the spectral ripple of iterated rippled noise (IRN) that is possibly resolved by the auditory system. Tonal signals were placed at spectral peaks and valleys of IRN maskers for a wide variety of IRN conditions that included different delays, number of iterations, and stimulus durations. The differences in the forward-masked thresholds of tones at spectral peaks and valleys were used to estimate spectral resolvability, and these results were compared to estimates obtained from a gamma-tone filter bank. The IRN spectrum has spectral peaks that are harmonics of the reciprocal of the delay used to generate IRN stimuli. As the number of iterations in the generation of IRN stimuli increases so does the difference in the spectral peak-to-valley ratio. For high number of iterations, long delays, and long durations evidence for spectral resolvability existed up to the 6th harmonic. For all other conditions spectral resolvability appeared to disappear at harmonics lower than the 6th, or was not measurable at all. These data will be discussed in terms of the role spectral resolvability might play in processing the pitch, pitch strength, and timbre of IRN stimuli. [Work supported by a grant from NIDCD.]

4:05

1pPP10. Searching for sources of variance in speech recognition: Young adults with normal hearing. Charles S. Watson and Gary R. Kidd (Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405)

In the present investigation, sensory-perceptual abilities of one thousand young adults with normal hearing are being evaluated with a range of auditory, visual, and cognitive measures. Four auditory measures were derived from factor-analytic analyses of previous studies with 18–20 speech and non-speech variables [G. R. Kidd *et al.*, *J. Acoust. Soc. Am.* **108**, 2641 (2000)]. Two measures of visual acuity are obtained to determine whether variation in sensory skills tends to exist primarily within or across sensory modalities. A working memory test, grade point average, and Scholastic Aptitude Test scores (Verbal and Quantitative) are also included. Preliminary multivariate analyses support previous studies of individual differences in auditory abilities (e.g., A. M. Surprenant and C. S. Watson, *J. Acoust. Soc. Am.* **110**, 2085–2095 (2001)) which found that

spectral and temporal resolving power obtained with pure tones and more complex unfamiliar stimuli have little or no correlation with measures of speech recognition under difficult listening conditions. The current findings show that visual acuity, working memory, and intellectual measures are also very poor predictors of speech recognition ability, supporting the independence of this processing skill. Remarkable performance by some exceptional listeners will be described. [Work supported by the Office of Naval Research, Award No. N000140310644.]

4:20

1pPP11. Temporal and spectral resolution of hearing in patients with precipitous hearing loss: Gap release of masking (GRM) and the role of cognitive function. Martin D. Vestergaard^{a)} (Tech. Univ. of Denmark, Ørsted-DTU Acoust. Technol., DK-2800 Lyngby, Denmark and Res. Ctr. Eriksholm, Kongevejen 243, DK-3070 Snekkerten, Denmark)

The purpose of this experiment was to measure temporal acuity and spectral resolution of hearing in new hearing-aid users over a period of time post-fitting, and to demonstrate the extent to which performance might change over time. For one-octave wide maskers with and without spectral and temporal gaps, masking was measured repeatedly over 3 months post-fitting. GRM was characterized as the release from masking under the gap conditions. The cognitive skills of the participants were assessed with two tests for measuring working memory capacity and lexical vigilance. The results showed that while the masking by one-octave wide noise maskers without any gaps was constant over time, GRM increased over time for maskers involving a temporal gap. Moreover, at low frequencies where the subjects had normal hearing-threshold levels, they performed as hearing-impaired for the spectral-gap condition. For the temporal-gap condition, they performed as normally hearing at both low and high frequencies. These results suggest that patients with precipitous hearing loss do not maintain normal spectral resolution through the low-frequency region, in which the hearing threshold levels are otherwise normal. Surprisingly, the results also showed moderate though highly significant correlation between lexical vigilance and GRM. [Work supported by the William Demant Foundation.] ^{a)}Currently at CNBH, Dept. Physiol., University of Cambridge, CB2 3EG Cambridge, UK.

4:35

1pPP12. Simulated phase-locking stimulation: An improved signal processing strategy for cochlear implant. Xihong Wu, Hongwei Qu, Jing Chen, Tianshu Qu (Nat. Key Lab. on Machine Percept. Speech and Hearing Res. Ctr., Dept. of Psych., Peking Univ., Beijing 100871, China), and Liang Li (Peking Univ., Beijing 100871, China)

Electrical stimulation of the auditory pathway produces different patterns of neural activity than those acoustically elicited. Traditional signal-processing strategies for cochlear implant usually do not utilize phase information contained in sound waves. Here, to evaluate potential advantages of introducing phase information to cochlear implant devices, a new signal processing method, so called simulated phase-locking stimulation (SPLS), was developed. To convey phase information of sound signals to the auditory nerve, electrical stimulation pulses were delivered at the zero-crossing time of sine waves of frequency bands after band-pass filtering and envelope extraction. The advantages of the SPLS method over the method of Continuous Interleaved Sampling (CIS+) were demonstrated by both objective evaluations, such as the spectro-temporal modulation index (STMI), and subjective evaluations, such as recognition of processed Chinese speech by normal hearing listeners under either noise (energetic) masking or speech (informational) masking conditions. The results suggest that the SPLS method is able to improve the function of cochlear devices by extracting and transferring fine-structure signals, which are important for cochlear-implant listeners to perceive tonal speech and music.

1p MON. PM

Session 1pSC**Speech Communication: Vowel Systems and Language Learners (Lecture/Poster Session)**

Linda Polka, Cochair

McGill Univ., School Communication Sciences and Disorders, 1266 Pine Ave., West, Montreal, QC H3G 1A8, Canada

Ocke-Schwen Bohn, Cochair

*English Dept., Aarhus Univ., Aarhus DK-8000, Denmark***Chair's Introduction—1:00*****Invited Papers*****1:05****1pSC1. Diversity of vowel systems.** Ian Maddieson (Dept of Linguist., Univ. of California, Berkeley, CA 94720)

Systems of vowels vary greatly across the world's languages while nonetheless conforming to certain general structural patterns. All languages have at least two qualitative distinctions between vowels based on the major parameters of height, backness and rounding, but probably none has more than 15 or so, and the modal number is 5. Generally these basic vowel qualities respect dispersion principles, but deviations can be considerable. When additional parameters, such as nasalization, length, phonation type and pharyngealization are included, the total number of vowel distinctions may easily exceed 40. These "additive" features never occur with a larger number of vowel qualities than those occurring in a "plain" series. Languages may differ markedly in the distributional patterns of their vowels as well as in their inventory. Some languages have different (usually reduced) vowel inventories in unstressed or other non-prominent positions; others constrain vowel sequences in (phonological) words through vowel harmony limitations. Co-occurrence patterns between vowels and consonants also vary greatly, as does the degree of coarticulation between vowels and neighboring segments. Learners must master all of these factors to speak an individual language fluently. Constraints that are universal or shared may be expected to facilitate this task.

1:30

1pSC2. Natural referent vowels guide the development of vowel perception. Linda Polka (School of Commun. Sci. and Disord., McGill Univ., Montreal, QC, Canada, linda.polka@mcgill.ca), Ocke-Schwen Bohn (Aarhus Univ., Denmark), and Monika Molnar (McGill Univ., Montreal, QC, Canada)

Certain vowels are favored across languages of the world. This selection bias has received a great deal of attention in linguistic theories seeking to explain vowel system typologies. In comparison, the role that specific vowels might play in the ontogeny of vowel perception has been more implicit. In this talk we will summarize recent findings that elucidate the functional significance of peripheral vowels in the development of vowel perception. Data from cross-language studies of infant vowel discrimination and vowel preference will be presented. This work shows that peripheral vowels have a perceptual priority for young infants and that this bias is independent of the phonemic status of the vowels presented in the perceptual task. Findings from cross-language experiments with adults reveal that language experience builds on the natural vowel biases observed in infancy. Adult data suggest that the natural bias remains in place in mature listeners unless the perceiver needs to override the bias to optimize perception of functional vowel differences. These findings support our proposal of a Natural Reference Vowel hypothesis as a framework for understanding the development of vowel perception and production. Specific avenues of research needed to elaborate this framework will be outlined. [Work supported by NSERC.]

1:55

1pSC3. Production-perception relationships during speech development. Lucie Menard (UQAM, Departement de linguistique, CP 8888, succ. Ctr.-Ville, Montreal, Canada H3C 3P8, menard.lucie@uqam.ca), Jean-Luc Schwartz, Louis-Jean Boe (Universite Stendhal, 38031 Grenoble Cedex 1, France), and Jerome Aubin (UQAM, CP 8888, succ. Ctr.-Ville, Montreal, Canada H3C 3P8)

It has been shown that nonuniform growth of the supraglottal cavities, motor control development, and perceptual refinement shape the vowel systems during speech development. In this talk, we propose to investigate the role of perceptual constraints as a guide to the speakers task from birth to adulthood. Simulations with an articulatory-to-acoustic model, acoustic analyses of natural vowels, and results of perceptual tests provide evidence that the production-perception relationships evolve with age. At the perceptual level, results show that (i) linear combination of spectral peaks are good predictors of vowel targets, and (ii) focalization, defined as an acoustic pattern with close neighboring formants [J.-L. Schwartz, L.-J. Boe, N. Vallee, and C. Abry, *J. Phonetics* **25**, 255–286 (1997)], is part of the speech task. At the production level, we propose that (i) frequently produced vowels in the baby's early sound inventory can in part be explained by perceptual templates, (ii) the achievement of these perceptual templates may require adaptive articulatory strategies for the child, compared with the adults, to cope with morphological differences. Results are discussed in the light of a perception for action control theory. [Work supported by the Social Sciences and Humanities Research Council of Canada.]

1pSC4. The influence of different native language systems on vowel discrimination and identification. Diane Kewley-Port (Dept. Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405, kewley@indiana.edu), Ocke-Schwen Bohn (Aarhus Univ., Denmark), and Kanae Nishi (Indiana Univ., Bloomington, IN 47405)

The ability to identify the vowel sounds of a language reliably is dependent on the ability to discriminate between vowels at a more sensory level. This study examined how the complexity of the vowel systems of three native languages (L1) influenced listeners perception of American English (AE) vowels. AE has a fairly complex vowel system with 11 monophthongs. In contrast, Japanese has only 5 spectrally different vowels, while Swedish has 9 and Danish has 12. Six listeners, with exposure of less than 4 months in English speaking environments, participated from each L1. Their performance in two tasks was compared to 6 AE listeners. As expected, there were large differences in a linguistic identification task using 4 confusable AE low vowels. Japanese listeners performed quite poorly compared to listeners with more complex L1 vowel systems. Thresholds for formant discrimination for the 3 groups were very similar to those of native AE listeners. Thus it appears that sensory abilities for discriminating vowels are only slightly affected by native vowel systems, and that vowel confusions occur at a more central, linguistic level. [Work supported by funding from NIHDCD-02229 and the American-Scandinavian Foundation.]

1pSC5. Cross-language comparisons of contextual variation in the production and perception of vowels. Winifred Strange (Ph.D. Program in Speech and Hearing Sci., CUNY-Grad. Ctr., 365 Fifth Ave., New York, NY 10016-4309, strangepin@aol.com)

In the last two decades, a considerable amount of research has investigated second-language (L2) learners problems with perception and production of non-native vowels. Most studies have been conducted using stimuli in which the vowels are produced and presented in simple, citation-form (lists) monosyllabic or disyllabic utterances. In my laboratory, we have investigated the spectral (static/dynamic formant patterns) and temporal (syllable duration) variation in vowel productions as a function of speech-style (list/sentence utterances), speaking rate (normal/rapid), sentence focus (narrow focus/post-focus) and phonetic context (voicing/place of surrounding consonants). Data will be presented for a set of languages that include large and small vowel inventories, stress-, syllable-, and mora-timed prosody, and that vary in the phonological/phonetic function of vowel length, diphthongization, and palatalization. Results show language-specific patterns of contextual variation that affect the cross-language acoustic similarity of vowels. Research on cross-language patterns of perceived phonetic similarity by naive listeners suggests that listener's knowledge of native language (L1) patterns of contextual variation influences their L1/L2 similarity judgments and subsequently, their discrimination of L2 contrasts. Implications of these findings for assessing L2 learners perception of vowels and for developing laboratory training procedures to improve L2 vowel perception will be discussed. [Work supported by NIDCD.]

3:10–3:20 Break

Contributed Papers

All posters will be on display and all authors will be at their posters from 3:20 p.m. to 4:30 p.m.

1pSC6. Perceptual assimilation and categorial discrimination of American vowels by Japanese listeners. Miwako Hisagi, Winifred Strange (Dept. of Speech and Hearing Sci., CUNY-Grad.-Ctr., 365 Fifth Ave., New York, NY 10016, mhisagi@hotmail.com), Reiko Akahane-Yamada, and Rieko Kubo (Adv. Telecommunications Res. Inst. Intl., 2-2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan)

Best's Perception Assimilation Model predicts that relative difficulty discriminating non-native (L2) contrasts is predictable from perceived similarity of L2 segments and native (L1) segments. Japanese listeners performed a categorial discrimination task in which 9 vowel pairs (6 adjacent height pairs, 3 front-back pairs) involving 6 tokens (2 speakers/3 repetitions) of each of 8 American vowels /i, i, e, æ, a, ʌ, u, u/ were tested in the context of hVba disyllables. In a second task, listeners were asked to categorize all stimuli with respect to which Japanese vowel they were most similar, and to rate their goodness on a 9-point Likert scale. Overall error rates on height pairs ranged from 1 percent to 29 percent, and on front/back pairs, from 1 percent to 18 percent. The most difficult height contrasts were /u-u/ and /a-ʌ/; perceptual assimilation patterns showed that these pairs were assimilated to the same Japanese vowels (Single Category or Category Goodness pattern) although /a-ʌ/ were assimilated to 2-mora versus 1-mora Japanese /a/, respectively. The most difficult front/back contrast was /æ-a/. Surprisingly, American /i-i/ was discriminated very well and were assimilated to different Japanese vowels /i, e/, respectively. In general, perceptual assimilation patterns predicted discrimination accuracy quite well. [Work supported by NIDCD.]

1pSC7. A signal detection theory-based analysis of American English vowel identification and production performance by native speakers of Japanese. Stephen Lambacher (Ctr. for Lang. Res., Univ. of Aizu, Aizu-Wakamatsu, Fukushima-ken, Japan 965-8580, steeve@u-aizu.ac.jp), William Martens (McGill Univ., Montreal, QC, Canada H3A 1E3), and Kazuhiko Kakehi (Chukyo Univ., Kaizu-cho, Toyota, 470-0393, Japan)

The identification and production performance by two groups of native Japanese of the American English (AE) vowels /æ/, /a/, /ʌ/, /ɔ/, /ɜ/ was measured before and after a six-week, identification training program. A signal detection theory (SDT) analysis of the confusion data, as measured by d' , revealed that all five AE vowels were more identifiable by the experimental trained group than the control untrained group. The d' results showed that /ʌ/ was less identifiable than /ɔ/ in the pretest, even though the percentage identification rate for /ʌ/ was slightly greater than that for /ɔ/. Both groups productions of a list of CVCs, each containing one of the target AE vowels, were presented to a group of native AE listeners in a series of identification tasks. The d' results revealed that the AE listeners could more sensitively identify the experimental groups post-test vowel productions than they could the control groups. SDT analysis also clarified an additional potentially confusing result: /ʌ/ was somewhat less identifiable than /ɜ/, despite the fact that the percentage identification rate for /ʌ/ was higher. Overall, the SDT-based analysis served to change the pattern of results observed for L2 vowel identification and influenced the interpretation of the data.

1pSC8. Training Japanese L2 learners to perceive difficult American vowel contrasts. Mieko Sperbeck (Dept. of Linguist., City Univ. of New York–Grad. School and Univ. Ctr., 365 Fifth Ave., New York, NY 10016-4309), Winifred Strange, and Kikuyo Ito (City Univ. of New York–Grad. School and Univ. Ctr., New York, NY 10016-4309)

This study compared two methods to improve perception of three vowels ae, ah, uh and examined generalization of training to new vowels eh, awe and to trained vowels by new speakers in new consonantal contexts and in real words. One training protocol began with hVba disyllables (sessions 1–4) and then introduced sentence-length utterances with vowels in gabVpa/gadVda trisyllables (sessions 5–9). The other protocol used sentence materials in all 9 sessions. Twenty-four participants were divided into the two training conditions; half of each group were relatively inexperienced while the others were more experienced. Identification pretests on the 5 vowels in sentence materials showed no differences in overall accuracy for inexperienced (51% correct) and experienced groups (54% correct). Post-tests showed that perception of trained vowels improved in both trained and non-trained trisyllable/sentence materials (65% correct for both training groups). There were no significant differences in amount of improvement for inexperienced and experienced participants. Training also led to improvement on trained vowels by new speakers in new contexts and to real words. However, performance on non-trained vowels did not improve (actually decreased for eh) suggesting a response bias in post-test identification performance. [Work Supported by NIH.]

1pSC9. Within- and across-language spectral and temporal variability of vowels in different phonetic and prosodic contexts: Russian and Japanese. Yana D. Gilichinskaya, Miwako Hisagi, Françoise F. Law II, Shari Berkowitz, and Kikuyo Ito (CUNY–Grad. Ctr., Dept. of Speech and Hearing Sci., 365 Fifth Ave., New York, NY 10016)

Contextual variability of vowels in three languages with large vowel inventories was examined previously. Here, variability of vowels in two languages with small inventories (Russian, Japanese) was explored. Vowels were produced by three female speakers of each language in four contexts: (Vba) disyllables and in 3-syllable nonsense words (gaC1VC2a) embedded within carrier sentences; contexts included bilabial stops (bVp) in normal rate sentences and alveolar stops (dVt) in both normal and rapid rate sentences. Dependent variables were syllable durations and formant frequencies at syllable midpoint. Results showed very little variation across consonant and rate conditions in formants for /i/ in both languages. Japanese short /u, o, a/ showed fronting ($F2$ increases) in alveolar context relative to labial context (1.3-2.0 Barks), which was more pronounced in rapid sentences. Fronting of Japanese long vowels was less pronounced (0.3 to 0.9 Barks). Japanese long/short vowel ratios varied with speaking style (syllables versus sentences) and speaking rate. All Russian vowels except /i/ were fronted in alveolar vs labial context (1.1-3.1 Barks) but showed little change in either spectrum or duration with speaking rate. Comparisons of these patterns of variability with American English, French and German vowel results will be discussed.

1pSC10. Reading skills and the discrimination of English vowel contrasts by bilingual Spanish/English-speaking children: Is there a correlation? Sandra Levey (Dept. of Speech-Lang.-Hearing Sci., Lehman College of the City Univ. of New York, 250 Bedford Park Blvd. W., Bronx, NY 10468)

This study examined the discrimination of English vowel contrasts in real and novel word-pairs by 21 children: 11 bilingual Spanish/English- and 10 monolingual English-speaking children, 8–12 years of age ($M = 10; 6$; $Mdn = 10; 4$). The goal was to determine if children with poor reading skills had difficulty with discrimination, an essential factor in reading abilities. A categorical discrimination task was used in an ABX discrimination paradigm: A (the first word in the sequence) and B (the second word in the sequence) were different stimuli, and X (the third word in the sequence) was identical to either A or to B. Stimuli were produced

by one of three different speakers. Seventy-two monosyllabic words were presented: 36 real English and 36 novel words. Vowels were those absent from the inventory of Spanish vowels. Discrimination accuracy for the English-speaking children with good reading skills was significantly greater than for the bilingual-speaking children with good or poor reading skills. Early age of acquisition and greater percentage of time devoted to communication in English played the greatest role in bilingual children's discrimination and reading skills. The adjacency of vowels in the $F1$ - $F2$ acoustic space presented the greatest difficulty.

1pSC11. English vowel learning by speakers of Mandarin. Ron I. Thomson (Dept. of Linguist., Univ. of Alberta, 4th Fl. Assiniboia Hall, Edmonton, AB, Canada T6G 2G5, rit@ualberta.ca)

One of the most influential models of second language (L2) speech perception and production [Flege, *Speech Perception and Linguistic Experience* (York, Baltimore, 1995) pp. 233–277] argues that during initial stages of L2 acquisition, perceptual categories sharing the same or nearly the same acoustic space as first language (L1) categories will be processed as members of that L1 category. Previous research has generally been limited to testing these claims on binary L2 contrasts, rather than larger portions of the perceptual space. This study examines the development of 10 English vowel categories by 20 Mandarin L1 learners of English. Imitation of English vowel stimuli by these learners, at 6 data collection points over the course of one year, were recorded. Using a statistical pattern recognition model, these productions were then assessed against native speaker norms. The degree to which the learners' perception/production shifted toward the target English vowels and the degree to which they matched L1 categories in ways predicted by theoretical models are discussed. The results of this experiment suggest that previous claims about perceptual assimilation of L2 categories to L1 categories may be too strong.

1pSC12. Vowel space development in a child acquiring English and Spanish from birth. Jean Andruski, Sahyang Kim (ASLP, Wayne State Univ., 581 Manoogian Hall, 906 W. Warren, Detroit, MI 48202), Geoffrey Nathan (Wayne State Univ., Detroit, MI 48202), Eugenia Casielles, and Richard Work (Wayne State Univ., Detroit, MI 48202)

To date, research on bilingual first language acquisition has tended to focus on the development of higher levels of language, with relatively few analyses of the acoustic characteristics of bilingual infants' and children's speech. Since monolingual infants begin to show perceptual divisions of vowel space that resemble adult native speakers' divisions by about 6 months of age [Kuhl *et al.*, *Science* **255**, 606–608 (1992)], bilingual children's vowel production may provide evidence of their awareness of language differences relatively early during language development. This paper will examine the development of vowel categories in a child whose mother is a native speaker of Castilian Spanish, and whose father is a native speaker of American English. Each parent speaks to the child only in her/his native language. For this study, recordings made at the ages of 2;5 and 2;10 were analyzed and $F1$ – $F2$ measurements were made of vowels from the stressed syllables of content words. The development of vowel space is compared across ages within each language, and across languages at each age. In addition, the child's productions are compared with the mother's and father's vocalic productions, which provide the predominant input in Spanish and English respectively.

1pSC13. Acoustic properties of vowels in clear and conversational speech by female non-native English speakers. Chi-Nin Li and Connie K. So (Dept. of Linguist., Simon Fraser Univ., Burnaby, BC, Canada V5A 1S6)

Studies have shown that talkers can improve the intelligibility of their speech when instructed to speak as if talking to a hearing-impaired person. The improvement of speech intelligibility is associated with specific

acoustic-phonetic changes: increases in vowel duration and fundamental frequency (F_0), a wider pitch range, and a shift in formant frequencies for F_1 and F_2 . Most previous studies of clear speech production have been conducted with native speakers; research with second language speakers is much less common. The present study examined the acoustic properties of non-native English vowels produced in a clear speaking style. Five female Cantonese speakers and a comparison group of English speakers were recorded producing four vowels (*i u ae a*) in /bVt/ context in conversational and clear speech. Vowel durations, F_0 , pitch range, and the first two formants for each of the four vowels were measured. Analyses revealed that for both groups of speakers, vowel durations, F_0 , pitch range, and F_1 spoken clearly were greater than those produced conversationally. However, F_2 was higher in conversational speech than in clear speech. The findings suggest that female non-native English speakers exhibit acoustic-phonetic patterns similar to those of native speakers when asked to produce English vowels clearly.

1pSC14. Greek perception and production of an English vowel contrast: A preliminary study. Václav J. Podlipský (Dept. of English and American Studies, Palacky Univ., Krizkovskeho 10, Olomouc, 77180, Czech Republic, vaclavjonaspodlipsky@centrum.cz)

This study focused on language-independent principles functioning in acquisition of second language (L2) contrasts. Specifically, it tested Bohn's Desensitization Hypothesis [in *Speech perception and linguistic experience: Issues in Cross Language Research*, edited by W. Strange (York Press, Baltimore, 1995)] which predicted that Greek speakers of English as an L2 would base their perceptual identification of English /i/ and /I/ on durational differences. Synthetic vowels differing orthogonally in duration and spectrum between the /i/ and /I/ endpoints served as stimuli for a forced-choice identification test. To assess L2 proficiency and to evaluate the possibility of cross-language category assimilation, productions of English /i/, /I/, and /e/ and of Greek /i/ and /e/ were elicited and analyzed acoustically. The L2 utterances were also rated for the degree of foreign accent. Two native speakers of Modern Greek with low and 2 with intermediate experience in English participated. Six native English (NE) listeners and 6 NE speakers tested in an earlier study constituted the control groups. Heterogeneous perceptual behavior was observed for the L2 subjects. It is concluded that until acquisition in completely naturalistic settings is tested, possible interference of formally induced meta-linguistic differentiation between a "short" and a "long" vowel cannot be eliminated.

1pSC15. Diphthongs in the repopulated vowel space. Anna Bogacka (Adam Mickiewicz Univ., Poznan, Poland and Bielefeld Univ., Germany, abogacka@ifa.amu.edu.pl)

The study examined 8 British English diphthongs produced by Polish learners of English, testing the diphthongs' quality, duration, nasalization, and occurrence of glottal stops before the diphthongs. There were twelve conditions in which the diphthongs were tested: word-initial, word-final, before a voiced obstruent, before a voiceless obstruent, before a nasal consonant, and before a nasal consonant followed by a fricative, and each of these conditions was tested in a stressed and unstressed position. The diphthongs were tested in real words, embedded in sentences, controlled for the stress position, rhythmic units, and length. The sentences were read by 8 female and 8 male Polish learners of English and control subjects. The aim of the phonetic analysis done with Praat, and employing the methodologies used by Flege (1995) for SLA and Peeters (1991) and Jacewicz, Fujimara, and Fox (2003) for diphthongs, is to examine the shape of the restructured vowel space (Liljencrants and Lindblom 1972; Stevens 1989). The approach taken here is termed Vowel Space Repopulation to emphasize that the vowel space of Polish speakers of English is re-structured by new categories in complex ways which are not adequately captured by traditional notions such as "transfer," "interference," or "interlanguage."

1pSC16. Training Japanese listeners to identify American English vowels. Kanae Nishi and Diane Kewley-Port (Dept. of Speech and Hearing Sci., Indiana Univ., 200 S. Jordan Ave., Bloomington, IN 47405)

Perception training of phonemes by second language (L2) learners has been studied primarily using consonant contrasts, where the number of contrasting sounds rarely exceeds five. In order to investigate the effects of stimulus sets, this training study used two conditions: 9 American English vowels covering the entire vowel space (9V), and 3 difficult vowels for problem-focused training (3V). Native speakers of Japanese were trained for nine days. To assess changes in performance due to training, a battery of perception and production tests were given pre- and post-training, as well as 3 months following training. The 9V trainees improved vowel perception on all vowels after training, on average by 23%. Their performance at the 3-month test was slightly worse than the posttest, but still better than the pretest. Transfer of training effect to stimuli spoken by new speakers was observed. Strong response bias observed in the pretest disappeared after the training. The preliminary results of the 3V trainees showed substantial improvement only on the trained vowels. The implications of this research for improved training of L2 learners to understand speech will be discussed. [Work supported by NIH-NIDCD DC-006313 & DC-02229.]

1pSC17. The development of vowel spaces in English- and Korean-learning infants' speech. Soyoung Lee (Univ. of Wisconsin at Milwaukee, P.O. Box 413, Milwaukee, WI 53201)

A previous study (Yang, 1996) revealed that the vowel spaces of adult speech differ between English and Korean. This study longitudinally investigated whether vowel spaces of English- and Korean-learning infants' speech demonstrated similar patterns to their ambient languages. Speech samples of English- and Korean-learning infants were collected at 12 and 24 months and transcribed by either native English- or Korean-speakers, respectively. First and second formants of each vowel were measured using LPC, spectral peak value, and spectrographic formant mid points. The vowel spaces between the two groups displayed similar patterns at 12 months although the frequency of occurrence of each vowel differed (e.g., [i] occurs more frequently in English than in Korean). However, the vowel spaces showed different patterns at 24 months. F_2 values for front vowels [i, e] were higher in English-learning infants' speech than those in Korean. [a] in Korean was located at a central position of vowel space while it was located at a back position in English. These patterns were similar to the adult vowel space of Korean and English. This study suggests that infants form vowel space similar to their own languages at around 24 months.

1pSC18. Spanish listeners' perceptual patterns for English /i/ and /I/. Geoffrey Stewart Morrison (Dept. of Linguist., Univ. of Alberta, Edmonton, AB, Canada, T6G 2E7, gsm2@ualberta.ca)

Spanish has five monophthongs which differ only in spectral properties. General Canadian English has ten monophthongs which differ in steady-state spectral properties and duration, and many nominal monophthongs have substantial diphthongization. Unlike Spanish, Canadian English also uses duration as a cue to postvocalic obstruent voicing. The present study investigates L1-English and L1-Spanish L2-English listeners perception of a Canadian English /bit, bit, bid, bid/ continuum varying in steady-state spectral values and duration. Several patterns emerge in the L1-Spanish listeners data, including a contrary pattern in which duration is used in the same direction as L1-English listeners but spectral properties in the opposite direction. With respect to /i/ and /I/ perception, these patterns are generally consistent with the stages of learning proposed by Escudero [Unpublished Masters Thesis, Edinburgh University, 2000]: L1-Spanish listeners cannot initially perceive the difference, next they use duration properties, later they begin to use spectral properties, and finally they have L1-English-like primary use of spectral properties and secondary use of duration. Production data adds additional insight into the relationship of the perceptual patterns to the stages of learning.

1pSC19. The acquisition of Taiwan Mandarin vowels by native American English speakers. Cyun-jhan Lin (NCTU Dept. of Foreign Lang. and Lit., Ta-Hsueh Rd. 1001, Hsinchu 300, Taiwan, moyal1980.fg92g@nctu.edu.tw)

Previous work on the production of English and French phones by native American English speakers indicated that equivalence classification prevent L2 learners from approximating L2 phonetic norms of similar phones and that learning French would not affect English speakers' production of L1 similar phone /u/ (Flege, 1987). In this study, there were five subjects, including 2 advanced native American English learners of Taiwan Mandarin, 2 basic native American English learners of Taiwan Mandarin, and 1 monolingual Taiwan Mandarin speaker. The corpus were 12 English words "heed, who'd, hod; leak, Luke, lock; beat, suit, bot; peat, suit, pot," and 12 Mandarin words [i, u, a; li, lu, la; pi, pu, pa; phi, phu, pha]. Both advanced and basic learners' production of English and Mandarin words and monolingual Taiwan Mandarin speaker's production of Mandarin words were directly recorded onto a PC. Vowel formants were taken from spectrograms generated by Praat. Preliminary results showed the vowel space of advanced learners between Taiwan Mandarin [i] and [u] was larger than that of basic learners, and closer to the Taiwan Mandarin norms. Besides, the vowel space between English [i] and [u] by basic learners was dramatically smaller than that of American English norms.

1pSC20. Evolution of the speech intelligibility of prelinguistically deaf children who received a cochlear implant. Marie-Eve Bouchard, Henri Cohen (Cognit. Neurosci. Ctr., Univ. of Quebec at Montreal, Montreal, QC, Canada, H3C 3P8), and Marie-Therese LeNormand (Robert Debre Hospital, Paris, France)

The 2 main objectives of this investigation are (1) to assess the evolution of the speech intelligibility of 12 prelinguistically deaf children implanted between 25 and 78 months of age and (2) to clarify the influence of the age at implantation on the intelligibility. Speech productions videorecorded at 6, 18 and 36 months following surgery during a standardized free play session. Selected syllables were then presented to 40 adults listeners who were asked to identify the vowels or the consonants they heard and to judge the quality of the segments. Perceived vowels were then located in the vocalic space whereas consonants were classified according to voicing, manner and place of articulation. 3 (Groups) \times 3 (Times) ANOVA with repeated measures revealed a clear influence of time as well as age at implantation on the acquisition patterns. Speech intelligibility of these implanted children tended to improve as their experience with the device increased. Based on these results, it is proposed that sensory restoration following cochlear implant served as a probe to develop articulatory strategies allowing them to reach the intended acoustico-perceptual target.

1pSC21. A comparison of vowel formant frequencies in the babbling of infants exposed to Canadian English and Canadian French. Karen Mattock, Susan Rvachew, Linda Polka, and Sara Turner (School of Commun. Sci. & Disord., McGill Univ., 1266 Pine Ave. West, Montreal, QC, H3G 1A8 Canada, karen.mattock@mail.mcgill.ca)

It is well established that normally developing infants typically enter the canonical babbling stage of production between 6 and 8 months of age. However, whether the linguistic environment affects babbling, either in terms of the phonetic inventory of vowels produced by infants [Oller & Eiler (1982)] or the acoustics of vowel formants [Boysson-Bardies *et al.*

(1989)] is controversial. The spontaneous speech of 42 Canadian English- and Canadian French-learning infants aged 8 to 11, 12 to 15 and 16 to 18 months of age was recorded and digitized to yield a total of 1253 vowels that were spectrally analyzed and statistically compared for differences in first and second formant frequencies. Language-specific influences on vowel acoustics were hypothesized. Preliminary results reveal changes in formant frequencies as a function of age and language background. There is evidence of decreases over age in the $F1$ values of French but not English infants vowels, and decreases over age in the $F2$ values of English but not French infants vowels. The notion of an age-related shift in infants attention to language-specific acoustic features and the implications of this for early vocal development as well as for the production of Canadian English and Canadian French vowels will be discussed.

1pSC22. Perception of vowels by learners of Spanish and English. Mariche Garcia-Bayonas (Univ. of North Carolina-Greensboro, 321 McIver Bldg., UNCG, Greensboro, NC 27402-5001, megarcia@uncg.edu)

This study investigates the perception of English vowels /i I/, /u U/, and /e E/ and Spanish /i u e/ by native-speakers (NS) and learners (L) and compares these two sets of vowels cross-linguistically. Research on the acquisition of vowels indicates that learners can improve their perception with exposure to the second language [Bohn and Flege (1990)]. Johnson, Flemming, and Wright (1993) investigated the hyperspace effect and how listeners tended to choose extreme vowel qualities in a method of adjustment (MOA) task. The theoretical framework of this study is Fleges (1995) Speech Learning Model. The research question is: Are vowels selected differently by NS and L using synthesized data? Spanish learners ($n=54$) and English learners ($n=17$) completed MOA tasks in which they were exposed to 330 synthetically produced vowels to analyze spectral differences in the acquisition of both sound systems, and how the learners vowel system may vary from that of the NS. In the MOA tasks they were asked to select which synthesized vowel sounds resembled the most the ones whose spelling was presented to them. The results include an overview of the vowel formant analysis performed, and which vowels are the most challenging ones to learners.

1pSC23. Perception of steady-state vowels and vowelless syllables by adults and children. Susan Nittrouer (Ctr. for Persons with Disabilities, Utah State Univ., 6840 Old Main Hill, Logan, UT 84322)

Vowels can be produced as long, isolated, and steady-state, but that is not how they are found in natural speech. Instead natural speech consists of almost continuously changing (i.e., dynamic) acoustic forms from which mature listeners recover underlying phonetic form. Some theories suggest that children need steady-state information to recognize vowels (and so learn vowel systems), even though that information is sparse in natural speech. The current study examined whether young children can recover vowel targets from dynamic forms, or whether they need steady-state information. Vowel recognition was measured for adults and children (3, 5, and 7 years) for natural productions of /dæd/, /dud/ /æ/, /u/ edited to make six stimulus sets: three dynamic (whole syllables; syllables with middle 50-percent replaced by cough; syllables with all but the first and last three pitch periods replaced by cough), and three steady-state (natural, isolated vowels; reiterated pitch periods from those vowels; reiterated pitch periods from the syllables). Adults scored nearly perfectly on all but first/last three pitch period stimuli. Children performed nearly perfectly only when the entire syllable was heard, and performed similarly (near 80%) for all other stimuli. Consequently, children need dynamic forms to perceive vowels; steady-state forms are not preferred.

4:30–5:00 Panel Discussion

Separate registration fee required to attend Tutorial lecture

MONDAY EVENING, 16 MAY 2005

REGENCY A AND B, 7:00 TO 9:00 P.M.

Session 1eID

Interdisciplinary: Tutorial Lecture: Automatic Speech Recognition

Fredericka Bell-Berti, Chair

St. John's Univ., Speech Communication Sciences and Theatre, 8000 Utopia Pkwy., Jamaica, NY 11439

Chair's Introduction—7:00

7:05

1eID1. Automatic speech recognition. Carol Espy-Wilson (Univ. of Maryland, Elect. & Comp. Eng. Dept., A. V. Williams Bldg., College Park, MD 20742)

Great strides have been made in the development of automatic speech recognition (ASR) technology over the past thirty years. Most of this effort has been centered around the extension and improvement of Hidden Markov Model (HMM) approaches to ASR. Current commercially-available and industry systems based on HMMs can perform well for certain situational tasks that restrict variability such as phone dialing or limited voice commands. However, the holy grail of ASR systems is performance comparable to humans—in other words, the ability to automatically transcribe unrestricted conversational speech spoken by an infinite number of speakers under varying acoustic environments. This goal is far from being reached. Key to the success of ASR is effective modeling of variability in the speech signal. This tutorial will review the basics of ASR and the various ways in which our current knowledge of speech production, speech perception and prosody can be exploited to improve robustness at every level of the system.

1p MON. PM

Session 2aAA

**Architectural Acoustics, Education in Acoustics, Noise and Psychological and Physiological Acoustics:
Topical Meeting on Classroom Acoustics—The Research Perspective II**

Murray R. Hodgson, Cochair

*Univ. of British Columbia, School of Occupational and Environmental Hygiene, 2206 East Mall,
Vancouver, BC V6T 1Z3, Canada*

Lily M. Wang, Cochair

Univ. of Nebraska-Lincoln, Architectural Engineering, Peter Kiewit Institute, 1110 South 67th St., Omaha, NE 68182-0681

Chair's Introduction—8:00

Invited Papers

8:05

2aAA1. Survey of conditions in New Zealand primary school classrooms and research on mechanisms influencing speech perception for children. George Dodd (School of Architecture, Univ. of Auckland, Private Bag 92019, Auckland, New Zealand, g.dodd@auckland.ac.nz)

A survey of 122 classrooms in New Zealand supports the need for low reverberation time and low background noise if classrooms are to be satisfactory. The RT of “poor” and “good” classrooms was found to be 0.6 s and 0.4 s respectively. Six “poor” classrooms were modified to reduce their RT to 0.4 s which changed them to “good” as judged by the users. The need for this low RT may result from children having a small integration time. A novel technique of reversed-segmented speech has verified that this is significantly smaller for children compared with adults. Activity noise in classrooms exhibits the cafe effect—a rising noise level as children compete to be heard—and we suggest this is caused by the Lombard effect. Values of the Lombard effect that we measured in a cohort of primary school children predict classroom levels similar to those observed. Present theory does not predict that reducing the RT from 0.6 to 0.4 s should significantly influence the cafe effect. Further work is planned to refine the theory and to identify if there are mechanisms not accounted for.

8:35

2aAA2. Auralization studies to develop a classroom questionnaire. Christian Nocke, Catja Hilge (Akustikbuero Oldenburg, Alte Raad 20a, D-26127 Oldenburg, Germany, info@akustikbuero-oldenburg.de), and Markus Meis (Hoerzentrum Oldenburg GmbH, D-26129 Oldenburg, Germany)

Modern, computer-based room acoustic modeling software allows apart from the calculation and visualization of sound fields also the auralization of room acoustic conditions. By these means the listening conditions in (virtual) class rooms can easily be modeled and modified. Furthermore, for each listening position a full set of room acoustical parameters is available. Corresponding measurements in real class rooms would require much effort; well-defined changes could only be reached by interior works in the building. So, computer modeling offers a larger variety of well-defined acoustic conditions. This contribution reports on the auralizations carried out for the development of a questionnaire for subjective evaluations in real class rooms. For this purpose six different virtual class rooms have been modeled. In each of the six situations auralizations at two different listening positions for male and female speakers have been deduced. The auralizations have been presented to 80 test persons. The objective room acoustic parameters have been correlated with the subjective judgements. The questionnaire has been used by Meis *et al.* in real life experiments on the subjective evaluation of listening conditions. Results of these experiments will be presented in another contribution to this session by M. Meis (2pAAa1).

8:55

2aAA3. Subjective and objective assessments in classrooms following acoustical renovation. Arianna Astolfi (Dept. of Energy Technologies, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129, Torino, Italy)

The effectiveness of an expensive acoustical intervention in an old Italian high school building has been assessed in this work. The school building has fifty classrooms, the majority of which were acoustically renovated. A subjective survey and measurements were performed in both the renovated and non renovated classrooms. With the competence of some psychologists from Turin University, a questionnaire was set up for the subjective analysis. The questionnaire, validated after numerous pilot tests, was submitted to all the students and the teachers in two periods of the year. The questions on acoustical features included questions on annoyance from room noise, reverberation, speech comprehension, overall acoustical satisfaction and the consequences of bad acoustical conditions. Apart from the acoustics, other aspects of environmental quality, such as the thermal and visual environmental features and IAQ were investigated. The statistical analysis of the subjective answers allowed aggregated information to be obtained on the users and different data to be correlated. The aim of the statistical correlation was to determine any significant relationships between the objective and subjective data, and between the overall satisfaction scores and the different environmental factors. The effects of bad environmental conditions and their influence on learning capacity were also examined.

9:15

2aAA4. Questionnaire study of perceived listening quality in renovated university classrooms. Murray Hodgson and Hugh Davies (UBC School of Occ. & Env. Hygiene, 3rd Fl., 2206 East Mall, Vancouver, BC, Canada V6T 1Z3)

The research reported here investigated the effect of renovations on student perception of the listening environment in university classrooms. It involved four large classrooms at the University of British Columbia which were studied before and after acoustical renovation. Details of the renovations were identified. Measurements of reverberation time, background-noise level and sound propagation were made in each case. Room-average Speech Transmission Indices were calculated as a physically-based summary quality measure, and their changes on renovation were determined. Questionnaires designed to determine student and instructor perception of the listening environment, as well as the personal, academic and environmental factors that might affect it, were administered to a large number of students taking classes in each classroom, and their instructors, before and after renovation. A summary Perceived Listening Ease (PLE) score, measuring student perception of the quality of the listening environment, was calculated for each questionnaire from the responses. Room-average PLE scores, and the changes on renovation, were then calculated. These results were related to the physical measurement results. Physical, personal, academic and environmental factors affecting the changes in PLE score were investigated. The implications of the results for classroom design were considered.

9:35

2aAA5. Acoustical conditions for speech communication in active elementary school classrooms. Hiroshi Sato (Inst. for Human Sci. & Biomed. Eng., Natl. Inst. of Adv. Industrial Sci. and Technol., 1-1-1 Higashi, Tsukuba, Japan) and John Bradley (Inst. for Res. in Construction Natl. Res. Council, Ottawa, Canada K1A 0R6)

Detailed acoustical measurements were made in 34 active elementary school classrooms with typical rectangular room shape in schools near Ottawa, Canada. There was an average of 21 students in classrooms. The measurements were made to obtain accurate indications of the acoustical quality of conditions for speech communication during actual teaching activities. Mean speech and noise levels were determined from the distribution of recorded sound levels and the average speech-to-noise ratio was 11 dBA. Measured mid-frequency reverberation times (RT) during the same occupied conditions varied from 0.3 to 0.6 s, and were a little less than for the unoccupied rooms. RT values were not related to noise levels. Octave band speech and noise levels, useful-to-detrimental ratios, and Speech Transmission Index values were also determined. Key results included: (1) The average vocal effort of teachers corresponded to louder than Pearsons Raised voice level; (2) teachers increase their voice level to overcome ambient noise; (3) effective speech levels can be enhanced by up to 5 dB by early reflection energy; and (4) student activity is seen to be the dominant noise source, increasing average noise levels by up to 10 dBA during teaching activities. [Work supported by CLLRnet.]

9:55

2aAA6. Investigation of noise issues at U-Hill Elementary School. Natalie York and Murray Hodgson (UBC Acoust. and Noise Res. Group, SOEH, 3rd Fl., 2206 East Mall, Vancouver, BC, Canada V6T 1Z3, nlyork@interchange.ubc.ca)

University Hill Elementary School (U-Hill) has a reputation amongst its teachers, staff and parents as being a noisy school. The present work is an assessment of the acoustical environment at U-Hill with a focus on evaluating how various noise sources contribute to the total noise levels measured in four classrooms. A questionnaire was completed by several of U-Hill's teachers to determine their reactions to the school's acoustical environments and what they perceived to be the prevalent noise sources. Standard physical measurements of classroom background noise and reverberation time were made and compared to existing acceptability criteria. Measurements of sound propagation and transmission were made to determine how noise propagates in the hallway, as well as into and between adjacent learning spaces, and the effect of open doors and door grates. Dosimeters were used to monitor sound levels at various hallway source locations and received in the four classrooms studied. Octave-band measurements of typical noise-source levels, such as children walking in the hallways and entering through school doors were made. The measurement data was used to estimate the contributions to total noise levels in the four classrooms of the various noise sources inside and outside the classrooms.

10:15–10:30 Break

10:30

2aAA7. Acoustical conditions of typical classrooms in Hong Kong. Kai Ming Li and Coriolanus C. L. Lam (Dept. of Mech. Eng., The Hong Kong Polytechnic Univ., Hung Hom, Hong Kong)

This paper presents measurement results of the acoustical environments of local schools in Hong Kong. In the measurements, several acoustical aspects that affect verbal communication in classrooms have been studied. These conditions include outdoor and indoor ambient noise levels, signal-to-noise ratios, reverberation time and the speech transmission index. Typical classrooms in many different schools and other higher-education institutions have been selected in the present study. Experimental results are compared with such national standards as USA (ANSI S 12.60 V 2002), Australian/New Zealand (AS/NZS 2107:2000), China (GB/T 15508 V 1995) and other national and industrial standards. This study will form the basis of devising acceptable standards for use in Hong Kong. [Work supported by the Research Grants Council of the SAR Government, the Research Committee of the Hong Kong Polytechnic University and Architectural Services Department of the Hong Kong SAR Government.]

10:50

2aAA8. Acoustic and social design of schools—ways to improve the school listening environment. Mechthild Hagen (Education Dept., Ludwig-Maximilians-Univ. of Munich, Leopoldstr. 13, D-80802 Muenchen, Germany, m_hagen@primedu.uni-muenchen.de)

Results of noise research indicate that communication, and as a result, teaching, learning and the social atmosphere are impeded by noise in schools. The development of strategies to reduce noise levels has often not been effective. A more promising approach seems to be to pro-actively support the ability to listen and to understand. The presentation describes the approach to an acoustic and social school design developed and explored within the project “GanzOhrSein” by the Education Department of the Ludwig-Maximilians-University of Munich. The scope includes an analysis of the current “school soundscape,” an introduction to the concept of the project to improve individual listening abilities and the conditions for listening, as well as practical examples and relevant research results. We conclude that an acoustic school design should combine acoustic changes in classrooms with educational activities to support listening at schools and thus contribute to improving individual learning conditions and to reducing stress on both pupils and teachers.

11:10

2aAA9. New schools design: Acoustics as main target. Luigi Maffei and Paola Lembo (Dispama, Faculty of Architecture, Second Univ. of Naples, Via S. Lorenzo, 81031 Aversa (Ce), Italy)

The effects of poor intelligibility and high background noise levels on the cognitive development of school children and on the dissatisfaction of teachers has been largely investigated. National standards have been implemented and attempts to harmonize these standards in international guidelines are ongoing. All these activities have led to the awareness that design of new schools must be centered on the achievement of a good acoustic environment. At this point a strong research effort to study and implement best solutions must be conducted, in collaboration, by architects, acousticians, pedagogues, psychologists, builders and acoustic materials producers. Recently an international competition for the planning of new primary schools in Rome, Italy has been announced. The aim of the competition is to study new architectural and running features of primary schools to obtain, among other parameters such as lighting, low cost energy solutions and air quality, the control of reverberation time, sound insulation and mechanical equipments noise. In these school buildings, as innovative requirement, children must be also able to elaborate interpretative hypothesis of physical phenomena such as sound emission and perception and be aware of their influence on these phenomena. Different possible solutions are presented.

Contributed Papers

11:30

2aAA10. Speech intelligibility metrics in small unoccupied classrooms. Matthew E. Cruikshank, Melinda J. Carney, and Dominique J. Cheenne (Dept. of Audio Arts & Acoust., Columbia College Chicago, Chicago, IL 60605)

Nine small volume classrooms in schools located in the Chicago suburbs were tested to quantify speech intelligibility at various seat locations. Several popular intelligibility metrics were investigated, including Speech Transmission Index (STI), %Alcons, Signal to Noise Ratios (SNR), and 80 ms Useful/Detrimental Ratios (U80). Incorrect STI values were experienced in high noise environments, while the U80s and the SNRs were found to be the most accurate methodologies. Test results are evaluated against the guidelines of ANSI S12.60-2002, and match the data from previous research.

11:45

2aAA11. The prediction of speech intelligibility in classrooms using computer models. Stephen Dance and Roger Dentoni (Dept. of Eng. Systems, London South Bank Univ., Borough Rd., London, UK)

Two classrooms were measured and modeled using the industry standard CATT model and the Web model CISM. Sound levels, reverberation times and speech intelligibility were predicted in these rooms using data for 7 octave bands. It was found that overall sound levels could be predicted to within 2 dB by both models. However, overall reverberation time was found to be accurately predicted by CATT 14% prediction error, but not by CISM, 41% prediction error. This compared to a 30% prediction error using classical theory. As for STI: CATT predicted within 11%, CISM to within 3% and Sabine to within 28% of the measured value. It should be noted that CISM took approximately 15 seconds to calculate, while CATT took 15 minutes. CISM is freely available on-line at www.whyverne.co.uk/acoustics/Pages/cism/cism.html

Session 2aAB

Animal Bioacoustics: Methodology for Measurements of Auditory Evoked Potentials in Aquatic Mammals I

Mardi C. Hastings, Chair

*Office of Naval Research, ONR 341, 800 North Quincy St., Arlington, VA 22217***Chair's Introduction—8:00****Invited Papers****8:05****2aAB1. Conditions for evoked-potential audiometry in odontocetes.** Alexander Ya. Supin (Inst. of Ecology and Evolution, Russian Acad. of Sci., 33 Leninsky Prospect, 119071 Moscow, Russia)

Currently auditory brainstem evoked potentials (ABR) become widely used for audiometry in odontocetes. Depending on the goal, the ABR technique differs with respect of (i) electrode positions, (ii) stimulus parameters, and (iii) threshold evaluation procedure. (I) Optimal electrode positions are vertex (for binaural responses), lateral (for monaural responses), or pharyngeal (for some express investigations). (II) The shorter the stimulus and the wider its spectrum, the more robust the response. Thereafter, narrow-band stimuli provoke low-amplitude ABR with a short dynamic range, but the goal of investigation often requires keeping the stimulus spectrum narrow. Sinusoidally amplitude modulated (SAM) sounds have narrow spectra, thereby provoking low amplitude envelope-following response (EFR); however, EFR can be extracted from noise by Fourier analysis. Modulation rate for SAM sounds must fit a spectrum peak of the ABR waveform. High-frequency sounds are more effective to provoke ABR and EFR than low-frequency ones, so ABR technique is better applicable for measurements in high-frequency ranges. (III) For precise threshold evaluation, low-amplitude responses must be extracted from noise. Apart from the averaging procedure, cross-correlation (for single ABR) or Fourier (for EFR) analysis is helpful to extract and measure low response amplitudes. [Work supported by Russian Basic Research Foundation, Russian President Grant, ONR.]

8:35**2aAB2. Auditory evoked potential measurement methodology for odontocetes and a comparison of measured thresholds with those obtained using psychophysical techniques.** Paul E. Nachtigall, Michelle Yuen, T. Aran Mooney, and Kristen Taylor (Marine Mammal Res. Program, Hawaii Inst. of Marine Biol., Univ. of Hawaii, Honolulu, HI 96822)

Most measurements of the hearing capabilities of toothed whales and dolphins have been taken using traditional psychophysical procedures in which the animals have been maintained in laboratory environments and trained to behaviorally report the sensation or difference of acoustic stimuli. Because of the advantage of rapid data collection, increased opportunities, and new methods, Auditory Evoked Potentials (AEPs) have become increasingly used to measure audition. The use of this new procedure calls to question the comparability of the established literature and the new results collected with AEPs. The results of behavioral and AEP methods have been directly compared with basic audiogram measurements and have been shown to produce similar (but not exactly the same) values when the envelope following response procedure has been used and the length of the stimulus is taken into account. The AEP methods allow possible audiometric opportunities beyond those available with conventional psychophysics including: (1) the measurement of stranded dolphins and whales that may never be kept in laboratories, (2) the testing of stranded animals for hearing deficits perhaps caused by overexposure to noise, and (3) passive testing of hearing mechanisms while animals actively echolocate. [Work supported by the Office of Naval Research and NOAA-NMFS.]

9:00**2aAB3. Auditory evoked potential measurements with cetaceans.** David Mann, Mandy Cook (USF College of Marine Sci. & Mote Marine Lab., 140 7th Ave. S., St. Petersburg, FL 33701, dmann@marine.usf.edu), Gordon Bauer (New College of Florida & Mote Marine Lab., Sarasota, FL 34243), Wendi Fellner (Walt Disney World, Lake Buena Vista, FL), and Randy Wells (Chicago Zoological Society and Mote Marine Lab., Sarasota, FL 34236)

Auditory evoked potentials (AEPs) allow researchers to measure the hearing abilities of animals that would be difficult or impossible to train for behavioral measurements of hearing. The hearing abilities of live-stranded cetaceans and wild dolphins can only be made with AEP techniques. In these situations, time with the animal is often restricted to an hour or less, and there is often little control over the acoustic environment in which the tests are performed. AEP measurements may be made while the animals are in air or in shallow pools. For cetaceans in air, sounds are typically presented with a suction cup jawphone. For cetaceans in water, sounds may be presented in a direct field (with the transducer located at some distance from the test subject) or with a jawphone. In each of these situations it is important to understand how thresholds derived from AEP measurements compare with behavioral hearing measurements. Examples of AEP measurements from wild and live-stranded cetaceans are presented to illustrate their usefulness and the constraints under which these measurements must be made. AEP measurements from bottlenose dolphins in air and in water are also compared with their behavioral audiograms.

9:25

2aAB4. Narrow band pulses as stimuli in an auditory brain stem recording study with a harbor porpoise. Kristian Beedholm and Lee A. Miller (Inst. of Biol., Univ. of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark, lee@biology.sdu.dk)

We have studied several aspects of hearing by a harbor porpoise using the ABR method with pulsed stimuli. Experiments were conducted on a male porpoise in collaboration with Fjord and Baelt, Kerteminde, Denmark. The animal had suction cups containing silver electrodes placed near the blowhole and near the dorsal fin. When fitted with the electrodes he moved to an underwater listening post where his outgoing sonar signal could be used to trigger a phantom echo. EEG signals were amplified differentially and averaged over a variable number of presentations depending on trial duration and experiment. For studying the frequency/intensity response, narrow band pulsed stimuli were generated and presented in several ways. One way was to use the impulse response of a B&K 1/3 octave filter bank (set to 80, 100, 125, or 160 kHz) as a stimulus. This stimulus was presented in both a passive hearing task, when a signal generator triggered the echo, and in an active experiment, where the echo was time locked to the animals emitted signal. Our results show the best response at 125 kHz and indicate a slight, but significantly higher response in the active mode. The latter has a methodological explanation. [Work supported by ONR.]

9:50

2aAB5. Auditory evoked potentials (AEP) methods for population-level assessment of hearing sensitivity in bottlenose dolphins. Dorian Houser (BIOMIMETICA, 7951 Shantung Dr., Santee, CA 92071) and James Finneran (Space and Naval Warfare Systems Ctr., San Diego, CA 92152)

A portable system for recording auditory evoked potentials (AEP) was developed to rapidly assess the hearing sensitivity of dolphins in air. The system utilizes a transducer embedded in a silicone suction cup to deliver amplitude modulated tones to the dolphin through the lower jaw. Frequencies tested range from 10–150 kHz and testing of both ears is completed within 90 min. AEP-determined thresholds from one subject were benchmarked against that subject's direct field behavioral audiogram to quantify variation between the two methods. To date, AEP audiograms have been obtained from over 30 bottlenose dolphins. Considerable individual variation in frequency-specific hearing sensitivity was observed. Some high-frequency hearing loss was observed in relatively young (early 20s) and old (35+ years) animals; conversely, age was not necessarily related to hearing loss as several animals greater than 40 years of age had good hearing sensitivity across the range of tested frequencies. Profound hearing loss typically occurred at higher frequencies. Decline in sensitivity was rapid in all cases and began between 50–60 kHz. Increased sample size of hearing sensitivity in dolphins suggest that the use of audiometric functions from single animals as representative of population level audiometry might be misleading.

10:15–10:30 Break

10:30

2aAB6. Electrophysiological methods for hearing assessment in pinnipeds. Colleen Reichmuth Kastak, David Kastak (Univ. of California Santa Cruz, Long Marine Lab., 100 Shaffer Rd., Santa Cruz, CA 95060, coll@ucsc.edu), James J. Finneran (SPAWARSSYSCEN San Diego, San Diego, CA 92152), Dorian S. Houser (BIOMIMETICA, Santee, CA 92071), and Alexander Supin (Inst. of Ecology and Evolution, 119071 Moscow, Russia)

Studies of auditory sensitivity in marine mammals generally rely on behavioral psychophysical methodologies. While these studies are the standard for hearing assessment in marine mammals, data are limited to only a few individuals representing a small proportion of species. Accumulating research on dolphin auditory physiology has resulted in the refinement of electrophysiological methods appropriate for odontocete cetaceans and an increase in available audiometric information. Electrophysiological methods have also been used with pinnipeds, but there are significant gaps in our understanding of pinniped auditory physiology that must be addressed before such approaches can be broadly applied to investigations of pinniped hearing. We are taking a bottom-up approach to developing suitable methods for evoked potential audiometry in pinnipeds, including technology transfer from studies of cetaceans and other mammals, mapping of response amplitude with respect to recording positions on the skull, characterization of responses in relationship to various stimulus types and presentation parameters, and determination of whether useful frequency-specific data can be reliably obtained using electrophysiological methods. This approach is being taken with representative pinniped species including California sea lions (*Zalophus californianus*), harbor seals (*Phoca vitulina*), and northern elephant seals (*Mirounga angustirostris*) using both training and chemical immobilization techniques. [Work supported by NOPP.]

10:55

2aAB7. Ground-truthing evoked potential measurements against behavioral conditioning in the goldfish, *Carassius auratus*. Randy J. Hill (College of Marine Sci., Univ. of South Florida, 140 7th Ave. S., St. Petersburg, FL 33701) and David A. Mann (Univ. of South Florida, St. Petersburg, FL 33701)

Auditory evoked potentials (AEPs) have become commonly used to measure hearing thresholds in fish. However, it is uncertain how well AEP thresholds match behavioral hearing thresholds and what effect variability in electrode placement has on AEPs. In the first experiment, the effect of electrode placement on AEPs was determined by simultaneously recording AEPs from four locations on each of 12 goldfish, *Carassius auratus*. In the second experiment, the hearing sensitivity of 12 goldfish was measured using both classical conditioning and AEP's in the same setup. For behavioral conditioning, the fish were trained to reduce their respiration rate in response to a 5 s sound presentation paired with a brief shock. A modified staircase method was used in which 20 reversals were completed for each frequency, and threshold levels were determined by averaging the last 12 reversals. Once the behavioral audiogram was completed, the AEP measurements were made without moving the fish. The recording electrode was located subdermally over the medulla, and was inserted prior to classical conditioning to minimize handling of animal. The same sound stimuli (pulsed tones) were presented and the resultant evoked potentials were recorded for 1000–6000 averages. AEP input–output functions were then compared to the behavioral audiogram to compare techniques for estimating behavioral thresholds from AEP data.

2aAB8. Electrophysiological responses to vibratory stimulation in the brainstem of the tadpole. Andrea M. Simmons, Aimee Two, and Seth S. Horowitz (Dept. Psych. and Neurosci., Brown Univ., Providence, RI 02912, Andrea_Simmons@brown.edu)

Maturation of saccular responses to vibratory stimuli in 3 axes was measured in the *Rana catesbeiana* tadpole across metamorphic development. The animal was suspended in a water-filled circular aluminum dish attached to a shaker table system consisting of 5 vibration exciters producing sinusoidal, translational movements in the *x*, *y*, and *z* planes. Multi-unit responses were recorded from the medulla with tungsten or platinum-iridium microelectrodes. Recording sites were electrolytically marked to verify electrode placement. Data were collected from 16 animals in late larval stages, before onset of metamorphic climax. Most recording sites were sensitive to vibrations in the *z* (vertical) direction, indicating a saccular origin of the response. Vibration frequencies of 40 Hz or below produced the strongest activity, with thresholds below a displacement of 10 nm. Secondary sensitivity was sometimes observed to either *x* (body-longitudinal) or *y* (body-lateral) directions of movement. Although some sites produced strong phase-locked activity to the stimulation, others showed more diverse temporal patterns of response. Active recording sites were found in the medial vestibular nucleus, lateral vestibular nucleus, reticular gray, and dorsal superior olivary nucleus. These data suggest that tadpole saccule is functional early in larval development, with sensitivity similar to that observed in adult bullfrogs.

TUESDAY MORNING, 17 MAY 2005

PLAZA A, 9:00 TO 11:45 A.M.

Session 2aAO

Acoustical Oceanography, Underwater Acoustics and Animal Bioacoustics: Acoustics and Ocean Observatories I

Bruce M. Howe, Chair

Applied Physics Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698

Chair's Introduction—9:00

Invited Papers

9:05

2aAO1. Initial results from the Pacific Ocean Shelf Tracking System: A permanent continental-scale acoustic tracking array for fisheries research. David W. Welch (Chief Scientist, Census of Marine Life Project POST, Kintama Res. Corp., 4737 Vista View Crescent, Nanaimo, BC, Canada V9V 1N8)

The Census of Marine Life is developing POST, a seabed acoustic array for tracking marine animals. The long-range plan involves the deployment of 30 or more cross-shelf monitoring lines forming a permanent continental-scale array, each consisting of autonomous seabed nodes spaced at roughly 1 km intervals. Nodes would be modular and use an acoustic modem to periodically communicate with an overhead ship, which would upload data and download new programming. 2004 was a large-scale test of the tracking array component using 135 km of listening lines and 1050 acoustically tagged juvenile salmon (12–15 cm long). Detection rates for individual fish crossing 20 km long acoustic lines were approximately 91%, and precise measurements of migration timing, speed of movement and survival were obtained for both the freshwater and early marine phases. In subsequent years we plan to include additional oceanographic sensors to provide detailed data on changes in bottom temperature, salinity, and currents over time and implement acoustic modems to allow remote data upload from a visiting boat. These data could then be meshed with the fish movement data to describe how animals move relative to changes in the three dimensional structure of the ocean that they are migrating through.

9:25

2aAO2. Autonomous underwater vehicle networks as integrated acoustic observation systems. Henrik Schmidt (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

Autonomous Underwater Vehicles (AUV) are rapidly being transitioned into operational systems for national defense, offshore exploration, and ocean science. However, the use the robotic platforms as components of integrated acoustic sensing systems is still at an early stage. Thus, for example, with recent advances in underwater navigation and communication, and collaborative robotics, networks of AUVs may have significant potential as moving tomographic networks for optimally and adaptively providing in-situ environmental estimates. Such concepts are important to the development and operation of environmenatally adaptive, autonomous littoral surveillance systems by providing environmental inversion specifically emphasizing the parameters significant for the actual platform configuration. Also such multi-AUV concepts have significant potential for subbottom characterization in the deep ocean, of increased importance in the offshore industry. For such systems to be operationally feasible, the limited underwater communication bandwidth makes it necessary for such systems to involve significant on-board processing, in turn requiring the acoustic sensing to be fully integrated with an onboard modeling capability and the platform navigation and control. The development of such new nested, distributed processing concepts, autonomously integrating sensing, modeling and control, is a significant challenge to the acoustics community. [Work supported by ONR.]

2aAO3. Communications performance of an undersea acoustic large-area network. Hannah A. Kriewaldt (Phys. Dept., Naval Postgrad. School, Monterey, CA 93943, hakesle@nps.edu) and Joseph A. Rice (SSC San Diego, Naval Postgrad. School, Monterey, CA 93943)

The U.S. Navy is developing Seaweb acoustic networking capability for integrating undersea systems. Seaweb architectures generally involve a wide-area network of fixed nodes consistent with future distributed autonomous sensors on the seafloor. Mobile nodes including autonomous undersea vehicles (AUVs) and submarines operate in the context of the grid by using the fixed nodes as both navigation reference points and communication access points. In October and November 2004, Theater Anti-Submarine Warfare Exercise (TASWEX04) showcased Seaweb in its first fleet appearance. This paper evaluates the TASWEX04 Seaweb performance in support of networked communications between a submarine and a surface ship. Considerations include physical-layer dependencies on the 9–14 kHz acoustic channel, such as refraction, wind-induced ambient noise, and submarine aspect angle. [Work supported by SSC San Diego.]

10:05–10:25 Break

10:25

2aAO4. Acoustic navigation for ultra-precise seafloor mapping. Donald M. Hussong and J. Grant Blackinton (Fugro Seafloor Surveys, Inc., 2727 Alaskan Way, Seattle, WA 98121)

Numerous autonomous underwater vehicles (AUVs) are operating with very precise swath bathymetric mapping systems in ocean depths to 3 000 m and greater. Although these systems acquire data from positions known to a few centimeters accuracy with respect to the vehicle position, the positioning of the AUVs is generally no better than tens of meters at great depths. Although not yet assembled into an operational system, the technology now exists to acoustically navigate an AUV to within a 10 cm in x , y , and z . This begins by establishing active reference positions on the seafloor, known to within a few centimeters absolutely, using acoustic mirror transponders that transmit and receive streams of pulse coded signals from an accurately navigated surface vessel. These bottom reference positions can then be the core of a long baseline acoustic navigation system which, when coupled with constant monitoring of oceanographic conditions to calculate pressure corrections for depth measurements, will provide 10-cm navigation accuracy for the mapping AUV. Implementation of this technology will enable observational mapping of highly seismic areas where frequent earthquakes can cause many meters of deformation of the seafloor, and can generate massive tsunamis.

Contributed Papers

10:45

2aAO5. NEPTUNE Canada—status and planning. Brian D. Bornhold (NEPTUNE Canada, P.O. Box 1700, STN CSC, Univ. of Victoria, Victoria, BC, Canada V8W 2Y2, bornhold@uvic.ca)

Stage 1 of the joint Canada-U.S. NEPTUNE seafloor observatory has been funded by the Canada Foundation for Innovation and the British Columbia Knowledge Development Fund with an overall budget of \$62.4 million. The network is designed to provide as close to real-time data and images as possible to be distributed to the research community, government agencies, educational institutions and the public via the Internet. Covering much of the northern segment of the Juan de Fuca Plate, this first phase of the NEPTUNE project is scheduled to be installed, with an initial suite of “community experiments”, in 2008. As part of the planning, NEPTUNE Canada held a series of three workshops to develop the science plans for these “community experiments”; these experiments have a budget of approximately \$13 million. The experiments will cover the gamut of oceanographic science themes including various aspects of: ocean climate and marine productivity, seabed environments and biological communities, fluids at ocean ridges, gas hydrates and fluids on continental margins, plate tectonics processes, associated earthquakes and tsunamis. The next three years will be spent developing and testing the necessary instrumentation for deployment on the network.

11:00

2aAO6. Ambient sound budgets. Jeffrey A. Nystuen and Bruce M. Howe (Appl. Phys. Lab., 1013 NE 40th St., Seattle, WA 98105)

The underwater ambient sound field contains a lot of information about geophysical, biological and anthropogenic activities in the ocean. Identification and quantification of these signals will complement more traditional measurements collected on future ocean observatories, including systems cabled to shore and remote mooring arrays. Long-term ambient sound measurements using a low duty cycle recorder are reported from remote deep ocean moorings, coastal shelf moorings, inland waterways and even a deep, quiet fresh water lake. Identifiable geophysical signals

include accurate measurements of wind speed and rainfall over a wide range of wind speeds and rainfall rates. The character and temporal patterns of biological activity, shipping and recreational boating are monitored. The overall ambient sound budgets will be reported for these different environments. Acoustic monitoring is available from sub-surface moorings, remote locations and under severe weather conditions where and when other ocean instrumentation is likely to fail. [Work supported by ONR Ocean Acoustics.]

11:15

2aAO7. The role of ocean observatories in monitoring for potential effects of man-made sound on the marine environment. David L. Bradley (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804), Gerald L. D’Spain (Scripps Inst. of Oceanogr., La Jolla, CA 93940-0701), James H. Miller (Univ. of Rhode Island, Narragansett, RI 02882), and George V. Frisk (Florida Atlantic Univ., Dania Beach, FL 33004-3023)

An important question is whether or not man-made sounds have an adverse long-term impact on the ocean environment. Clear evidence exists that an impact has occurred on some individual animals in a few cases. However, according to the 2003 National Academy of Sciences report, the degree of impact on any marine mammal population or on ocean ecosystems as a whole is unknown. An essential component of an approach to addressing this question is an ocean noise/marine ecosystem monitoring system. The monitoring regions should be global in extent and include biologically sensitive areas. The effort should be sustained so that long-term trends in ocean noise levels can be evaluated. An important aspect of the data collection effort is the type and quality of ancillary information. Ancillary data on the marine ecosystems allows determination of any impacts and data on the sources of marine noise and marine environmental properties which are required to develop metrics for evaluating and predicting the characteristics of the noise field. The ocean observatory system presently being contemplated could play an important role in establishing

an ocean noise monitoring capability. [Work supported through the National Ocean Partnership Program, with sponsorship from ONR, NOAA, NSF, and USGS.]

11:30

2aAO8. A nearshore observatory for Antarctic krill: Analysis of the spatial variability in their distribution and abundance. Joseph D. Warren (Southampton College, 239 Montauk Hwy, Southampton, NY 11968, joe.warren@liu.edu) and David A. Demer (Southwest Fisheries Sci. Ctr., La Jolla, CA 92037)

Annual surveys of the distribution and abundance of Antarctic krill are conducted by the United States Antarctic Marine Living Resources Program in order to assist fishery managers in conserving this economic and

ecologic resource. These surveys utilize a large vessel which does not sample close (within 10 km) to the South Shetland Islands. These islands are home to many krill predators (fur seals and penguins) who often forage in these nearshore waters. In order to better understand the abundance and distribution of the krill in these waters and to determine the physical and biological factors affecting the krill, a nearshore observatory has been created that combines multiple sensors and platforms to provide improved spatial and temporal coverage of the nearshore region of Livingston Island. This study utilizes a large research vessel, a small inflatable vessel, an array of instrumented buoys, and a pilot project involving an Autonomous Underwater Vehicle. This presentation will discuss acoustic backscatter measurements made from both large and small vessels. Results from the most recent field season (Jan–Feb 2005) will be presented as well as data from previous years (2002, 2004). [Work supported by the NSF Office of Polar Programs and NOAA.]

TUESDAY MORNING, 17 MAY 2005

REGENCY A, 8:00 TO 11:50 A.M.

Session 2aBB

Biomedical Ultrasound/Bioresponse to Vibration and Physical Acoustics: Diagnostic and Therapeutic Ultrasound Relevant to the Brain

Pierre D. Mourad, Cochair

Applied Physics Lab., Univ. of Washington, Neurosurgery, 1013 NE 40th St., Seattle, WA 98195

E. Carr Everbach, Cochair

Dept. of Engineering, Swarthmore College, 500 College Ave., Swarthmore, PA 19081-1397

Invited Papers

8:00

2aBB1. Opening the blood-brain barrier with ultrasound. Pierre Mourad (APL/Neurosurgery, Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, pierre@apl.washington.edu)

The blood-brain barrier (BBB) constitutes a unique anatomical structure within the brain that prevents many therapeutic chemicals within the blood from entering the brain. Research at the University of Washington has shown that it is possible to transiently disrupt the BBB without damage in an intra-operative setting, as verified by the flux of vital dyes from the blood-stream into the brain. This talk will review the research we have done to date on this subject, and discuss potential applications of this approach to BBB disruption.

8:20

2aBB2. Feasibility of noninvasive ultrasound delivery for tumor ablation and targeted drug delivery in the brain. Kullervo Hynynen, Nathan McDannold, Greg Clement, Jason White (Dept. of Radiol., Brigham and Women's Hospital, 75 Francis St., Boston, MA 02115), Lisa Treat, Xiangtao Yin, Ferenc Jolesz, Nikolai Sheikov (Brigham and Women's Hospital, Boston, MA 02115), and Natalia Vykhodtseva (Brigham and Women's Hospital, Boston, MA 02115)

The objective of our research during the past few years has been to develop multichannel ultrasound phased arrays for noninvasive brain interventions. We have been successful in developing methods for correcting the skull induced beam distortions and thus, are able to produce sharp focusing through human skulls. This method is now being tested for thermal ablation of tumors, with results from animal studies demonstrating feasibility. In addition, the ability of ultrasound to open the blood-brain barrier (BBB) locally has been explored in animal models. The results suggest that the transcranial ultrasound exposures can induce BBB opening such that therapeutic agents can be localized in the brain. This tool is especially powerful since the beam can be guided by MR images, thus providing anatomical or functional targeting. This talk will review our current status in this research, which ultimately aims for the clinical use of this methodology.

8:40

2aBB3. Transcranial power M-mode Doppler ultrasound for diagnosis of patent foramen ovale. Mark Moehring and Merrill Spencer (Spencer Technologies, 701 16th Ave., Seattle, WA 98122, mm@spencertechnologies.com)

Patent foramen ovale (PFO) is a right-to-left shunt (RLS) which communicates blood from the right to left atrium of the heart. PFO has been associated with stroke and, more recently, with migraine headache. Diagnosis of RLS can be accomplished effectively with transcranial power M-mode Doppler ultrasound (PMD). PMD is a modality which can be performed without the sedation

required by the more invasive diagnostic technique using transesophageal echocardiography. PMD for this application consists of 2 MHz pulse Doppler ultrasound with placement of sample gates at 2 mm intervals along the single-transducer beam axis, and 8 kHz pulse repetition rate (PMD100M, Spencer Technologies). Doppler power versus depth is constructed every 4ms, using 33 sample gates. Bubble microemboli injected in the venous system and moving across a PFO present as high intensity tracks on a PMD image, as emboli transit from the heart to the brain and through the observed cerebral vasculature. Use of PMD in this context has been reported in the clinical literature [M. P. Spencer, M. A. Moehring, J. Jesurum *et al*, *J. Neuroimaging* **14**, 342–349 (2004)]. This talk surveys the basic technical features of PMD for sensing PFO-related showers of bubble microemboli, and how these features provide clues to the severity of PFO.

Contributed Papers

9:00

9:15

2aBB4. Frequency-dependent ultrasound transmission through the human skull. P. Jason White, G. T. Clement, and Kullervo Hynynen (Dept. of Radiol., Brigham and Women's Hospital, Harvard Med. School, Boston, MA 02115)

The development of large-aperture multiple-source transducer arrays for ultrasound transmission through the human skull has demonstrated the possibility of controlled acoustic energy delivery into the brain parenchyma. The individual control of acoustic parameters from each ultrasound source allows for the correction of distortions arising from transmission through the skull bone and also opens up the possibility for electronic steering of the acoustic focus within the brain. To determine the efficacy of frequency modulation with such a device, this study examines the frequency dependence of ultrasound transmission in the range of 0.6–1.4 MHz through a series of seven points on each of three *ex vivo* human calvaria. Using broadband pulses, it is shown that the reflected spectra from the skull reveal information regarding the transmission energies at specific frequencies. In fact, there exists an inverse correlation between the reflected pressure amplitude and the transmitted pressure amplitude such that, for the frequency range under examination, approximately three local minima in the reflection spectra with the corresponding three peaks in transmission are observed. A noninvasive determination of the most efficient transmission frequency for a localized spatial region on a skull can yield an improvement of up to 3.5-dB in ultrasound intensity transmission.

2aBB5. Numerical simulation of ultrasound thermotherapy of brain with a scanned focus transducer. Sohrab Behnia, Farzan Ghalichi, Amin Jafari, and Ashkan Bonabi (Dept. of physics, IAU, Univ. of Ourmi, Iran, s.behnia@iaurmia.ac.ir)

Brain tumors are one of the most difficult ones to treat. The margin between destruction of the tumor and damage to the surrounding tissue is narrow in the brain. Ultrasound could be an effective treatment because of its ability to propagate deep in tissue and induce temperature rise at the focus while leaving the surrounding tissue intact. This study investigates whether using a fix-focus transducer could destroy brain tumor cells, in a cost effective manner which reduces the treatment time significantly. In this work an appropriate fix-focus transducer was designed considering effective parameters and limitations which are dominant in this case. Then a real 2-D brain model was constructed from a MR image. A piece of the skull bone has been removed to allow ultrasound to propagate into the brain. The resultant pressure field and the temperature rise were calculated by Rayleigh integral and bio-heat equation on the model. The obtained results were promising indicating that toxic temperatures could be obtained in short treatment times. This could be of great advantage especially in treating primary brain tumors.

Invited Papers

9:30

2aBB6. Ultrasound-accelerated thrombolysis using microbubbles. William Culp (Interventional Radiol., Univ. of Arkansas for Medical Sci., Little Rock, AR 72205-7199, CulpWilliamC@uams.edu)

Current thrombolytic therapy for ischemic stroke reaches less than 3% of cases and is only moderately successful. Numerous studies have demonstrated increased thrombolytic activity when ultrasound is delivered to clot in the presence of tissue plasminogen activator. Initial human reports using trans-cranial Doppler technology (2 MHz) to deliver continuous ultrasound to intracranial clot are very promising with improved clot lysis and no increase in symptomatic bleeding. However, another human study using low-frequency therapeutic ultrasound resulted in excessive bleeding and was discontinued. Microbubble augmented ultrasound clot lysis has proven successful in several studies ranging from lysis of very small peripheral clots in rabbits to very large clots in dialysis grafts in dogs. Early human studies of thrombosed dialysis grafts show success and no adverse events. Intracranial clot lysis in pigs has been successful with both intra-arterial and intravenous microbubble techniques. Better definition of possible combination therapies and of efficacy and safety is still required in animal models, but human studies should be designed to avoid some of the complications of current therapy. This presentation will review the subject, present recent advances, and define some requirements for successful clot destruction in peripheral vessels and intracranial vessels.

9:50

2aBB7. Ultrasound enhanced thrombolysis: Clinical evidence. Andrei V. Alexandrov (The Univ. of Texas–Houston Med. School, MSB 7.044, 6431 Fannin St., Houston, TX 77030)

Phase II CLOTBUST randomized clinical trial (Houston, Barcelona, Edmonton, Calgary) evaluated patients with acute ischemic stroke due to intracranial occlusion and treated with intravenous tissue plasminogen activator (TPA) within 3 h of symptom onset. Randomization: monitoring with pulsed wave 2 MHz transcranial Doppler (TCD) (Target) or placebo monitoring (Control). Safety: symptomatic bleeding to the brain (sICH). Primary end-point: complete recanalization on TCD or dramatic clinical recovery by the total NIHSS score <3, or improvement by >10 NIHSS points within 2 hours after TPA bolus. All projected 126 patients were randomized 1:1 to target (median NIHSS 16) or control (NIHSS 17). sICH: 4.8% Target, 4.8% Controls. Primary end-point was achieved by 31 (49%, Target) versus 19 (30%, Control), $p < 0.03$. At 3 months, 22 (42% Target) and 14 (29% Control) patients achieved favorable outcomes. Continuous TCD monitoring of intracranial occlusion safely augments TPA-induced arterial recanalization, and 2 MHz diagnostic ultrasound has a positive biological activity that aids systemic thrombolytic therapy. For the first time in clinical medicine, the CLOTBUST trial provides the evidence that ultrasound enhances thrombolytic activity of a drug in humans thereby confirming intense multi-disciplinary experimental research conducted worldwide for the past 30 years.

10:30

2aBB8. The effect of 40 kHz ultrasound on tissue plasminogen activator-induced clot lysis in three *in vitro* models. Marlien Pieters, Rob T. Hekkenberg, Marrie Barrett-Bergshoeff, and Dingeman C. Rijken (TNO Prevention and Health, Leiden, Netherlands)

In previous work from the same laboratory, high-frequency ultrasound (US) (3 MHz) was shown to promote *in vitro* fibrinolysis through enhanced supply of plasminogen to the clot surface. The application of high-frequency US is limited *in vivo* due to tissue heating. Low-frequency US, however, has less tissue heating and improved penetration. Internal plasma clot lysis and external lysis with compacted and non-compacted plasma clots were used to determine the magnitude of the effect of low-frequency US (40 kHz; 0.5 W/cm²) on tissue plasminogen activator-induced lysis and to elucidate the mechanisms behind the effect. Ultrasound enhanced lysis in all three models, with the largest effects (4-fold) in the external lysis model with compacted plasminogen-poor clots. The acceleration effect of ultrasound in this model decreased with increasing *t*-PA—and decreasing plasminogen concentrations. Ultrasound had a much smaller effect in this model when compacted plasminogen-rich clots were used. In the external lysis, non-compacted clot model, ultrasound resulted in consistently higher lysis rates. The acceleration effect of lysis, increased slightly (1.3 to 1.8-fold) with increasing *t*-PA—and decreasing plasminogen concentrations. Plasminogen supply to the clot surface was again shown to be an important contributor to ultrasound-enhanced lysis. [M. Pieters, R. T. Hekkenberg, M. Barrett-Bergshoeff, and D. C. Rijken. *Ultrasound in Med & Biol* **30**, 1545–1552 (2004).]

Contributed Papers

10:50

2aBB9. Confocal microscopy movies of fibrin clots during ultrasound-accelerated thrombolysis. E. Carr Everbach (Swarthmore College, Swarthmore, PA 19081, ceverba1@swarthmore.edu), Irina N. Chernysh, and John W. Weisel (Univ. of Penn School of Med, Philadelphia, PA 19104)

Blood clots made of human purified fibrin (white clots) were insonified with 1 MHz pulsed ultrasound during observation by fluorescence confocal microscopy. A deconvolution microscope allowed extremely thin sheets (0.2 μm) of fibrin to be viewed at a resolution of 0.2 μm per pixel, and the clot microstructure visualized. Acoustic pressure amplitudes from 0.1 to 0.8 MPa (peak-to-peak) were inferred using the image blur of 0.6- μm -diameter polystyrene spheres coated with FITC fluorescent label present in the clots. Acoustic pulse widths of 1 ms and pulse repetition frequencies of 125 Hz reduced clot heating to less than 3°C during each 30-minute exposure. Still 100 μm by 100 μm images were recorded every 10 seconds during pauses in insonification, to produce time-lapse movies that are compared with movies made during sham ultrasound exposures.

11:05

2aBB10. Acousto-mechanical and thermal properties of clotted blood. Volodymyr M. Nahirnyak (Dept. of Phys., Univ. of Cincinnati, Cincinnati, OH 45221-0511), S. Wang Yoon, and Christy K. Holland (Univ. of Cincinnati, Cincinnati, OH 45267-0586)

The efficacy of ultrasound-assisted thrombolysis as an adjunct treatment of ischemic stroke is being widely investigated. In order to determine the role of ultrasound hyperthermia in the process of blood clot disruption, the thermal and acousto-mechanical properties of clotted blood were measured *in vitro*. Whole blood clots were prepared from either fresh porcine or human blood by aliquoting 1.5 or 2.0 ml into 10 ml glass tubes (BD Vacutainer™, Franklin Lakes, NJ), immersing the tubes in a 37°C water bath for three hours and storing the clots at 5°C for at least three days prior to assessment of the properties, which ensured complete clot retraction. Direct calorimetric measurements using calibrated *E*-type thermocouples (Omega Engineering, Inc., Stamford, CT) were performed to determine the heat capacity and thermal conductivity of the human and porcine thrombi against a standard fluid, saline [0.9%]. The amplitude coefficient of attenuation of the clots was determined from 120 kHz to 3.5 MHz with a calibrated hydrophone (TC4038, RESON, Inc., Goleta, CA) in a 20±2°C water bath using the substitution method. The experimentally measured values of heat capacity, density, and thermal conductivity of porcine clotted blood are 3.23±0.46 J/g·K, 1.058±0.014 g/cm³, and 0.52±0.14 W/m·K. The attenuation coefficient ranged from 0.10 to 0.30 Nepers/cm over 120 kHz to 3.5 MHz. Measurements of the acousto-

mechanical and thermal properties of clotted blood can be helpful in theoretical modeling of ultrasound hyperthermia in ultrasound-assisted thrombolysis.

11:20

2aBB11. Measurements of the effect of haemolysis on speed of sound and attenuation through suspensions of red blood cells at 15 MHz. Nicola Z. Hobbs and Constantin-C. Coussios (Dept. of Eng. Sci., Univ. of Oxford, Parks Rd., Oxford OX1 3PJ, U.K.)

When blood flows through an artificial environment for prolonged periods of time, the red blood cells are gradually destroyed due to the high shear stresses to which they are exposed. The ability to measure the degree of haemolysis accurately and in real time is becoming increasingly important, especially in the context of organ preservation and of the implantation of heart-assist devices as a long-term method for treating heart failure. A recent study indicated that the backscattering coefficient (BSC) of red cell suspensions containing damaged red blood cells is significantly higher than the BSC of suspensions containing exclusively healthy cells. Using a specially designed acoustic chamber, suspensions containing different proportions of healthy and damaged cells were exposed to 15-MHz ultrasound. A thin tungsten wire immersed in each suspension was used to quantify the speed of sound and attenuation through increasingly haemolyzed suspensions. It is intended to utilize a combination of measurements of speed of sound, attenuation and backscattering in order to solve the inverse problem and to develop an on-line, non-invasive device for quantifying the degree of haemolysis in artificial circuits.

11:35

2aBB12. Lytic efficacy of apoli protein E2 (ApoE2) and recombinant tissue plasminogen activator (rt-PA) treatment with 120 kHz ultrasound in an *in-vitro* human clot model. Jason M. Meunier, Jason Y. Cheng (Dept. of Emergency Medicine, Univ. of Cincinnati, 231 Albert Sabin Way, Cincinnati, OH 45267, meuniejn@uc.edu), Joseph F. Clark (Univ. of Cincinnati Medical Ctr., Cincinnati, OH 45267), and George J. Shaw (Univ. of Cincinnati, Cincinnati, OH 45267)

Currently, the only FDA approved therapy for acute ischemic stroke is recombinant tissue plasminogen activator (rt-PA). However rt-PA has substantial side effects such as hemorrhage. This has led to interest in other potential therapies. For example, ultrasound (US) increases the lytic efficacy of rt-PA. Also, apolipoprotein E2 (ApoE2) increases rt-PA activity. This suggests combining US, ApoE2 and rt-PA to improve thrombolysis, but the efficacy is not known. Here, the lytic efficacy of apoE2, rt-PA and 120 kHz US is measured in a human clot model. Whole blood was obtained from volunteers, after local institutional approval. Clots were formed in 1.7 mm micropipettes, and placed in a water tank that allowed microscopic video imaging during US and thrombolytic exposure. Clots

were treated with rt-PA ($[rt-PA]=3.15 \mu\text{g/ml}$), rt-PA and apoE2 ($[apoE2]=9.8 \mu\text{g/ml}$), or rt-PA, apoE2 and 120 kHz US (0.35 MPa, PRF=1667 Hz, 80% duty cycle) for 15 min at 37°C in human plasma. Clot lysis was visually recorded and the lysis depth (LD) determined from these data using an image analysis algorithm. LD was linear with time for all treatments ($R^2 \geq 0.81$), allowing the determination of a lytic rate (LR).

LR was found to be 0.35 ± 0.03 , 1.55 ± 0.11 , and $0.75 \pm 0.04 \mu\text{m/min}$ for the rt-PA, rt-PA and apoE2, and US treated groups respectively. The thrombolytic efficacy of rt-PA is enhanced by ApoE2. The interaction of 120 kHz with apoE2 and rt-PA showed a reduced lytic efficacy compared with rt-PA and apoE2 treatment alone. It is possible that US interferes with the ApoE2-mediated activation of rt-PA.

TUESDAY MORNING, 17 MAY 2005

PLAZA B, 10:00 A.M. TO 12:00 NOON

Session 2aED

Education in Acoustics: Hands-On Experiments for High School Students

Uwe J. Hansen, Chair

Physics Dept., Indiana Univ., Terre Haute, IN 47809 10:00

Students and senior scientists will staff approximately 20 experiment stations for local high school students who will perform the experiments. These students will thus be exposed to a variety of acoustics principles and get some practical hands-on experience with a number of research tools. Regular ASA meeting participants are welcome to the session, as long as their participation does not interfere with student hands-on activities.

TUESDAY MORNING, 17 MAY 2005

BALMORAL, 8:30 A.M. TO 12:05 P.M.

Session 2aMU

Musical Acoustics: Virtual Musical Instruments

Julius O. Smith, Cochair

4360 Miller Ave., Palo Alto, CA 94306-4410

Tamara Smyth, Cochair

Simon Fraser Univ., School of Computer Science, Surrey, BC V3T 5X3, Canada

Invited Papers

8:30

2aMU1. A physical model and experimental testbed for real-time simulation of flute-like instruments. Patricio de la Cuadra (Ctr. for computer research in music and acoustic, CCRMA, Stanford Univ., CA 94305-8180), Benoît Fabre (Univ. Paris 6, 75252 Paris Cedex 05, France), Jonathan S. Abel (Universal Audio, Inc., Santa Cruz, CA 95060), and Julius O. Smith III (Stanford Univ., CA 94305-8180)

In the last decades, a wealth of experimental data has been gathered concerning sound production in flute-like instruments. These data have driven the development and fine tuning of analytical models for the physical processes involved. Here, a number of these results are integrated into a real-time physical model of a flute-like instrument. The model proposed is one-dimensional and driven by a single input pressure. The system was designed in the spirit of a laboratory testbed, where geometric and performance parameters can be controlled in real time, and in which important physical quantities can be viewed in real time, either in the time domain or frequency domain. In this way, the user may observe and validate the behavior of variables that would otherwise be hidden in the feedback loop. Additionally, the system was designed in a modular fashion, so as to allow independent substitution and adjustment of any of the model blocks. So as to allow real-time implementation and parameter control, a novel time-varying jet filter has been developed which not only closely matches measured jet oscillation amplitudes as a function of jet geometry and frequency, but may be rapidly slewed in response to quickly changing performance parameters.

8:55

2aMU2. A unified digital waveguide (infra)structure for synthesizing wind instrument sounds. Gary Scavone (Music Technol., Faculty of Music, McGill Univ., 555 Sherbrooke St. West, Montreal, QC, Canada H3A 1E3)

Virtual acoustic models provide researchers a means for testing and verifying their scientific understanding of physical systems. As well, they allow explorations in a world free from physical constraints. A digital waveguide structure is presented which implements a generalized air column, or resonator, model for wind instrument sound synthesis. This system has a physical inspiration and interpretation, a “blowed string,” but is an otherwise impossible reality. The structure produces harmonic resonant modes at integer multiples of a fundamental, with controls to vary harmonic content. It is useful for simulating both cylindrical and conical air columns of such instruments as clarinets, saxophones, flutes, oboes, and trumpets. Further, this structure robustly interconnects with several different non-linear “reed” models to produce a rich variety of wind instrument sounds. The result is a highly efficient computational algorithm that can be used to perform real-time sound synthesis on consumer-grade computers. [Work supported by the Canadian Foundation for Innovation.]

9:20

2aMU3. Complete geometric computer simulation of a classical guitar. Rolf Bader (Inst. of Musicology, Neue Rabenstr. 13, 20354 Hamburg, Germany)

The aim of formulating a complete model of a classical guitar body as a transient-time geometry is to get detailed insight into the vibrating and coupling behavior of the time-dependent guitar system. Here, especially the evolution of the guitars initial transient can be looked at with great detail and the produced sounds from this computer implementation can be listened to. Therefore, a stand-alone software was developed to build, calculate, and visualize the guitar. The model splits the guitar body into top plate, back plate, ribs, neck, inclosed air, and strings and couples these parts together including the coupling of bending waves and in-plane waves of these plates to serve for a better understanding of the coupling between the guitar parts and between these two kinds of waves. The resulting waveforms are integrated over the geometry and the resulting sounds show up the different roles and contributions of the different guitar body parts to the guitar sound. Here cooperation with guitar makers is established, as changes on the guitars geometry on the resulting sound can be considered as computer simulation and promising new sound qualities can then be used again in real instrument production.

9:45

2aMU4. An allpass filter design method with application to piano string synthesis. Jonathan S. Abel (Universal Audio, Inc., Santa Cruz, CA 95060), Julius O. Smith III (Stanford Univ., Stanford, CA 94305), and Julien Bensa (Universite Pierre et Marie Curie, Paris, France)

A nonparametric allpass filter design method for matching a desired group delay as a function of frequency is presented. The technique is useful in physical modeling synthesis of musical instruments exhibiting dispersive wave propagation in which different frequency bands travel at different speeds. While current group delay filter design methods suffer from numerical difficulties except at low filter orders, the technique presented here is numerically robust, producing an allpass filter in cascaded biquad form, and with the filter poles following a smooth loop within the unit circle. The technique was inspired by the observation that a pole-zero pair arranged in allpass form has 2π total group delay when integrated around the unit circle, regardless of the pole location. To match a given group delay characteristic, the method divides the frequency axis into sections containing 2π total group delay, and assigns a pole-zero allpass pair to each. In this way, the method incorporates an order selection technique, and by adding a pure delay to the desired group delay, allows the trading of increased filter order for improved fit to the frequency-dependent group delay. Results are presented for modeling the group delay of a stiff piano string under several computational constraints.

10:10–10:20 Break

10:20

2aMU5. Music and science meet at the micro level: Time-frequency methods and granular synthesis. Barry Truax (School of Commun., Simon Fraser Univ., Burnaby, BC, Canada V5A 1S6)

Musical research over the last century has become increasingly entwined with the areas of acoustics, psychoacoustics, and electroacoustics. One of the most striking results has been to push the frontiers of models of sound and music to the micro level, what is generally termed microsound. At this level, concepts of frequency and time are conjoined by a quantum relationship, with an uncertainty principle relating them that is precisely analogous to the more famous uncertainty principle of quantum physics. A class of methods of sound synthesis and signal processing known as time-frequency models have their basis at this quantum level such that changes in a signal's time domain result in spectral alterations and vice versa. One such method, granular synthesis and the granulation of sampled sound, produces results by the generation of high densities of acoustical quanta called grains. Such a radical shift has profound implications for not only our models of sound design, but also for the compositional methods that emerge as well as the role of the composer in guiding complex processes. The paper will argue that these models are examples of a class of complex systems exhibiting emergent form that create a new form of virtual music instrument.

10:45

2aMU6. Control of diffusion in digital waveguide reverberators. Patty Huang (CCRMA, Dept. of Music, Stanford Univ., Stanford, CA 94305)

Variable diffusion has been incorporated into many room acoustics models, which allows for the representation of materials with different scattering properties. By simulating the scattering of sound waves off surfaces in an acoustical enclosure, a more realistic sound field is produced. Due to constraints on the algorithm architecture and a preference for late reverberation which approaches decaying gaussian noise, it is less common for diffusion to be featured as an adjustable parameter in artificial reverberators. However, parameterized diffusion can be a tool for tailoring the character and density of the reverberant tail, enabling the generation of a greater variety of reverbs. This paper will present methods for implementing and controlling diffusion in artificial reverberation models based on digital waveguides. Sound examples will be given during the presentation.

11:10

2aMU7. Acoustics of the Intonarumori. Stefania Serafin (Medialogy, Aalborg Univ. Copenhagen, Lautrupvang 15, 2750 Ballerup, Denmark)

The Intonarumori were a family of musical instruments invented by the Italian futurist composer and painter Luigi Russolo. Each Intonarumori was made of a wooden parallelepiped sound box, inside which a wheel of different sizes and materials was setting into vibration a catgut or metal string. The pitch of the string was varied by using a lever, while the speed of the wheel was controlled by the performer using a crank. At one end of the string there was a drumhead that transmitted vibrations to the speaker. Unfortunately, all the original Intonarumori were destroyed after a fire during World War II. Since then, researchers have tried to understand the sound production mechanism of such instruments, especially by consulting the patents compiled by Russolo or by reading his book "The art of noise." In this paper we describe the acoustics of the Intonarumori. Based on such description, we propose physical models that simulate such instruments. The intonarumori's string is modeled using a one dimensional waveguide, which is excited either by an impact or a friction model. The body of the instrument is modeled using a 3-D rectangular mesh, while the horn is considered as an omnidirectional radiator.

Contributed Papers

11:35

2aMU8. Guidophone: A handheld virtual music instrument combining vocal tract geometry and hand gestures. Ryan Cassidy, Rodrigo Segnini, and Yi-Wen Liu (Ctr. for Comput. Res. in Music and Acoust., Stanford Univ., Stanford, CA 94305)

Vocal sounds provide an intuitive and appealing basis for virtual instruments. The former aspect allows most people to engage in a musical activity simply by repetition of a heard sound. Physical models in speech synthesis attempt, among other objectives, to approximate the geometry of the vocal tract required to produce specific sounds. Cook has developed one such model [Perry Cook, Identification of Control Parameters in an Articulatory Vocal Tract Model with Applications to the Synthesis of Singing, Stanford University, 1990], where the vocal tract is divided into tube sections that govern the transmission and reflection of acoustic energy at the junctions between sections. Tube section radii provide the model parameters. This paper explores the feasibility of using hand gestures to control those parameters. To that end, the problem of mapping hand gestures to the parameters of Cook's vocal synthesis model, with attention paid to transitions between successive phonemes, is explored. The physical basis of the model is reviewed. Next the relationship between the geometries implied by Cook's tract model and those determined by experiment is investigated. A means for mapping a number of parameters less than the degrees of freedom inherent in the model is presented. Finally, the details of a virtual instrument, controlled by a mechanical device recently developed for the work, are presented.

11:50

2aMU9. SCUBA: The Self-Contained Unified Bass Augmenter. Juan Pablo Cceres, Gautham J. Mysore, and Jeffrey Trevio (Ctr. for Comput. Res. in Music and Acoust., Stanford Univ., The Knoll, 660 Lomita, Stanford, CA 94305)

The Self-Contained Unified Bass Augmenter (SCUBA) is a new augmentative OSC (Open Sound Control) controller for the tuba. SCUBA allows new expressive possibilities by adding onboard continuous and discrete sensors to provide user-controlled parametric data for the processing of the instruments natural sound in Pd. SCUBA endows acoustic instruments with parametric control of virtual instruments and provides a means of integrating the output of a virtual instrument with the natural sound of the tuba. The user sends controller data from the augmented tuba interface to the virtual instrument via FSRs (Force-Sensitive Resistors) and buttons mounted on the existing instrument interface. An AVRMini microcontroller converts raw sensor data to OSC (Open Sound Control) messages, which are mapped in Pd to control virtual instrument parameters. Virtual instrument output is integrated into the solo instrument interface via satellite speakers mounted in the bell; this allows for mixing of the instruments natural sound with that of the virtual instrument to create the impression of a single instrument. This integration is the goal of the SCUBA project: by providing a flexible but unified control interface and acoustic output, traditional acoustic instrument interfaces can be augmented and paired with virtual musical instruments. [We would like to thank Michael Gurevich, Max Matthews, Bill Verplank, Pascal Stang, and classmates from Music 250 (Fall 04) for assistance in the realization of this project.]

Session 2aNS**Noise and Physical Acoustics: Urban Sound Propagation**

Jian Kang, Cochair

Univ. of Sheffield, School of Architecture, Western Bank, Sheffield S10 2TN, United Kingdom

Ning Xiang, Cochair

*Rensselaer Polytechnic Inst., Architecture, 110 8th St., Troy, NY 12180***Chair's Introduction—8:00*****Invited Papers*****8:05****2aNS1. Experimental study of sound propagation in a street.** Judicaël Picaut (Section acoustique routière et urbaine, LCPC, route de Bouaye, BP 4129, 44341 Bouguenais Cedex, France)

This paper presents an experimental study of the sound propagation in a street, performed during July 2002, in order to measure impulse responses at numerous locations in a street canyon, and, for several positions of the sound source. A specific attention has been paid on the processing of the experimental data, in order to obtain accurate values of the reverberation time and the steady-state sound pressure level, including also the compensation of the atmospheric attenuation. In total, 11592 useful data have been collected during this experiment, and are available by contacting the author. Reverberation times and sound levels, in the narrow street, are then analyzed and compared, for the 4 source positions in the street. This experimental study gives interesting results on the behavior of the sound field in the street.

8:30**2aNS2. A scale model study of parallel urban street canyons.** Maarten Hornikx, Jens Forssen, and Wolfgang Kropp (Dept. of Appl. Acoust., Chalmers Univ. of Technol., S-412 96 Göteborg, Sweden)

The access to quiet areas in cities is of increasing importance. Recently, the equivalent sources method for a two dimensional situation of parallel urban street canyons has been developed. One canyon represents a busy road, whereas the other is one without traffic; the quiet side. With the model, the transfer function between the two canyons can be calculated, as well as the influence of diffusion, absorption, and atmospheric turbulence on the transfer function. A scale model study of two parallel canyons has now been executed. A scale of 1:40 has been chosen and the maximum length sequence technique has been applied using the MLSSA system. Results of the scale model study have been compared to calculations with the equivalent sources method. The difference between a two-dimensional and a three-dimensional quiet side, between a coherent and an incoherent line source and the influence of absorption and diffusion has been investigated. The scale model study also gives insight in the evolution of the sound field in the time domain. [Work supported by the Swedish Foundation for Strategic Environmental Research (MISTRA).]

8:55**2aNS3. The effect on sound propagation in streets of the distribution of acoustic scatterers on facades.** David J. Oldham and Mostafa R. Ismail (School of Architecture, Univ. of Liverpool, Liverpool, L69 3BX, UK, djoldham@liv.ac.uk)

Until recently street noise propagation models assumed that faade reflections were either totally specular or totally diffuse. However, the characteristics of building facades are such that some energy will be reflected specularly and some diffusely. Many current models, therefore, use a scattering coefficient to characterize the diffuse reflection characteristics and the assumption is made that a fraction of the sound energy falling on any point on the facade will be scattered to the same degree. From observation it can be seen that facades are not homogeneous but consist of surfaces which are planar in nature but feature discontinuities such as windows, doors, ledges, etc. which will tend to act as distributed scattering centers. The effect of distributions of scattering centers on the propagation of sound in streets is examined using a combined image source and randomly scattered sound rays. Computer generation of street models using an object based approach enables the study of many different configurations using RAYNOISE software. Results are compared with those obtained assuming a homogeneous distribution of facade scattering using both RAYNOISE and a technique proposed by the authors. The results of a scale model experiment to investigate the effect of distributed scatters are presented.

9:20**2aNS4. Recent advances in modeling the propagation noise in high-rise cities.** Kai Ming Li (Dept. Mech., The Hong Kong Polytechnic Univ., Hung Hom, Hong Kong)

In the past few decades, we have witnessed a rapid growth in mechanized transport and transportation systems. We live in a transport-dominated society which has led to a marked improvement in dispersal of land use and to the increased opportunity for the separate development of residential, commercial, and industrial areas. In dense and high-rise cities, various modes of land transpor-

tation are the primary source of noise. The problem of transportation noise is not confined by political or social frontiers. It affects the rich who may live in a quiet residential area but who must make full use of transport to maintain their affluent existence, as well as the less fortunate who must live close to a highway, a major road, or an elevated railway line. A systematic development of the capability for accurate predictions of the propagation of land transportation noise in dense high-rise cities is highly desirable. This paper summarizes the current models for predicting sound fields in urban environments and gives an overview of the recent advances of various numerical models to predict the sound field in urban environments. [Work supported by the Research Grants Council of the Hong Kong SAR Government and the Hong Kong Polytechnic University.]

9:45–10:00 Break

10:00

2aNS5. Acoustic pulse propagation in an urban environment. Donald G. Albert (USA ERDC-CRREL, 72 Lyme Rd., Hanover, NH 03755-1290) and Lanbo Liu (USA ERDC-CRREL, Hanover, NH 03755-1290)

Experimental measurements conducted in a full-scale artificial village show that complex signatures are formed by multiple reflections and diffractions from buildings along the propagation path. A two-dimensional finite difference time domain (FDTD) simulation running on a personal computer allows this wave interaction to be studied in detail. Time reversal processing to locate a sound source in an urban area is investigated using this simulation method. The results demonstrate that as few as three non-line-of-sight sensors are sufficient to determine the source location, and that errors on the order of a meter in the building or sensor locations still allow the correct source location to be determined. [Work supported by U.S. Army.]

10:25

2aNS6. Sound field simulation and acoustic animation in urban squares. Jian Kang and Yan Meng (School of Architecture, Univ. of Sheffield, Western Bank, Sheffield S10 2TN, UK, j.kang@sheffield.ac.uk)

Urban squares are important components of cities, and the acoustic environment is important for their usability. While models and formulae for predicting the sound field in urban squares are important for their soundscape design and improvement, acoustic animation tools would be of great importance for designers as well as for public participation process, given that below a certain sound level, the soundscape evaluation depends mainly on the type of sounds rather than the loudness. This paper first briefly introduces acoustic simulation models developed for urban squares, as well as empirical formulae derived from a series of simulation. It then presents an acoustic animation tool currently being developed. In urban squares there are multiple dynamic sound sources, so that the computation time becomes a main concern. Nevertheless, the requirements for acoustic animation in urban squares are relatively low compared to auditoria. As a result, it is important to simplify the simulation process and algorithms. Based on a series of subjective tests in a virtual reality environment with various simulation parameters, a fast simulation method with acceptable accuracy has been explored. [Work supported by the European Commission.]

10:50

2aNS7. Modeling the characteristics of wheel/rail rolling noise. Wai Keung Lui, Kai Ming Li (Dept. of Mech. Eng., The Hong Kong Polytechnic Univ., Hung Hom, Hong Kong), and Glenn H. Frommer (MTR Corp. Ltd., Kowloon Bay, Hong Kong)

To study the sound radiation characteristics of a passing train, four sets of noise measurements for different train operational conditions have been conducted at three different sites, including ballast tracks at grade and railway on a concrete viaduct. The time histories computed by the horizontal radiation models were compared with the measured noise profiles. The measured sound exposure levels are used to deduce the vertical directivity pattern for different railway systems. It is found that the vertical directivity of different railway systems shows a rather similar pattern. The vertical directivity of train noise is shown to increase up to about $30\times$ before reducing to a minimum at 90° . A multipole expansion model is proposed to account for the vertical radiation directivity of the train noise. An empirical formula, which has been derived, compares well with the experimental data. The empirical model is found to be applicable to different train/rail systems at train speeds ranging up to 120 km/h in this study. [Work supported by MTR Corporation Ltd., Innovation Technology Commission of the HKSAR Government and The Hong Kong Polytechnic University.]

Contributed Papers

11:15

2aNS8. Microscopic traffic modelling in urban noise assessment. Bert De Coensel, Dick Botteldooren, Tom De Muer (Acoust. Group, Dept. of Information Technol., Ghent Univ., St. Pietersnieuwstraat 41, B-9000 Ghent, Belgium), Bert Peeters, and Gijsjan van Blokland (M+P Raadgevende Ingenieurs, NL-5260 CB Vught, The Netherlands)

The temporal structure of the urban soundscape can be rather complex, due to the presence of many screening and reflecting surfaces and many different sound sources, of which traffic noise is the most dominant. From the point of view of soundscape research, the background level as well as the time structure of noise peaks are important. However, these indicators

cannot be estimated easily by current noise prediction models, based on static traffic flows. Therefore, a dynamic traffic noise model was used, based on a microscopic traffic simulation. This way, individual vehicles can be traced, each having an associated set of noise sources that can depend on vehicle properties such as speed and acceleration, as well as on road properties such as the surface type. The model further consists of an ISO 9613 based propagation component, which can account for multiple reflections and diffractions. Maps of statistical noise levels, but also of more complicated measures reflecting the time structure of the soundscape, can be produced. This way, the soundscape contribution of single vehicles can be traced, as well as the influence of more general vehicle properties, such as the contribution of vehicle acceleration noise to the soundscape at junctions.

11:30

2aNS9. Increased damping in irregular resonators. Bernard Sapoval (Laboratoire de Physique de la Matière Condensée, Ecole Polytechnique, 91128 Palaiseau, France), Mark Asch (Université de Picardie Jules Verne, 80039 Amiens, France), Simon Felix, and Marcel Filoche (Ecole Polytechnique, 91128 Palaiseau, France)

The relation between shape and damping of shallow acoustical cavities has been studied numerically in the case where the dissipation occurs only on the cavity walls. It is first found that whatever the type of geometrical

irregularity, many, but not all the modes are localized. It is shown that the localization mechanism is what is called weak localization. The more irregular, the smaller the quality factors are found. However this effect is very different for the non-localized and the localized modes. For non-localized modes the damping increases roughly proportionally to the cavity surface. The localized modes are even more damped. These results generalize the results already obtained both numerically and experimentally on prefractal acoustical cavities. [B. Sapoval, O. Haeberle, and S. Russ, *J. Acoust. Soc. Am.* **102**, 2014–2019 (1997); B. Hebert, B. Sapoval, and S. Russ, *ibid.* **105**, 1567–1576 (1999)].

TUESDAY MORNING, 17 MAY 2005

REGENCY B, 7:55 A.M. TO 12:00 NOON

Session 2aPA

Physical Acoustics, Engineering Acoustics and Noise: Infrasound: A New Frontier in Monitoring the Earth I

Wayne N. Edwards, Cochair

Dept. of Earth Sciences, Univ. of Western Ontario, 1151 Richmond St., London, ON N6A 5B7, Canada

David E. Norris, Cochair

BBN Technologies, 1300 North 17th St., Arlington, VA 22209

Chair's Introduction—7:55

Invited Papers

8:00

2aPA1. Infrasound associated with the 2004 Sumatra megathrust earthquake and tsunami. Milton Garces, Pierre Caron, and Claus Hetzer (Infrasound Lab., Univ. of Hawaii, Manoa, 73-4460 Queen Kaahumanu Hwy., #119, Kailua-Kona, HI 96740-2638)

Multiple infrasound arrays in the Pacific and Indian Oceans that are part of the International Monitoring System (IMS) observed three distinct waveform signatures associated with the December 26, 2004 Sumatra earthquake and tsunami. Infrasound station in Palau, Diego Garcia, Madagascar, and Kenya observed (1) seismic arrivals (P , S and surface) from the earthquake, (2) T -phases, propagated along SOFAR channel in the ocean, and coupled back to the ground, and (3) infrasonic arrivals associated with either the tsunami generation mechanism or the motion of the ground above sea level. All signals were recorded by the pressure sensors in the arrays. The seismic and T -phase recordings are due to the sensitivity of the MB2000 microbarometers to ground vibration, whereas the infrasound arrivals correspond to dispersed acoustic waves propagated through atmospheric waveguides. It appears that the arrival of the tsunami, as well as oceanic infragravity waves following the tsunami, were not observed by the infrasound stations. We show the prominent features of the arrivals, present source location estimates, discuss the absence of a signal associated with the tsunami arrival, and consider the potential for using infrasound as a discriminant for tsunami genesis.

8:20

2aPA2. Acoustic surveillance for hazardous eruptions (ASHE). David McCormack (Geological Survey of Canada, 7 Observatory Crescent, Ottawa, ON, Canada K1A 0Y3, cormack@seismo.nrcan.gc.ca), Henry Bass (Univ. of Mississippi, University, MS 38677), Milton Garces (Univ. of Hawaii, Manoa, Kailua-Kona, HI 96740-2638), and Hugo Yepes (Escuela Politcnica Nacional, Casilla 17-01-2759, Quito, Ecuador)

The potential of using infrasound to rapidly identify explosive volcanic eruptions has been discussed in the environmental acoustics and aviation safety communities for some time. The ability of sounds in the 0.01–10 Hz range to propagate for long distances with little attenuation suggests broad-area regional monitoring with a modest number of observing sites is possible. The ASHE experiment tests both the practical utility of infrasound as a regional-scale volcanic eruption detection tool, and the feasibility of using such an infrasound system to provide timely operational alerts to aviation through Volcanic Ash Advisory Centres (VAACs). Several infrasound arrays are deployed in a volcanic region, sending data in real time to a central detector, and onward to participating VAACs for comparison with existing warning systems. The ASHE experiment will determine if infrasound can complement both seismic and satellite observations to improve monitoring of volcanic hazard. Continuous acoustic surveillance can reduce the ambiguity between eruptive and purely seismic activity in an active volcano and provide additional estimates for the onset time of an eruption. The onset time estimates can be used as triggers for ash transport models.

8:40

2aPA3. Estimation of bolide energy using observed infrasound signals. Wayne N. Edwards (Dept. of Earth Sci., Univ. of Western Ontario, London, ON, Canada N6A 5B7), Peter G. Brown (Univ. of Western Ontario, London, ON, Canada N6A 3K7), and Douglas O. ReVelle (Los Alamos Natl. Lab., Los Alamos, New Mexico, 87545)

The acoustic amplitude-yield relationships including formal errors for a population of large and well-observed (greater than 0.05 kton) bolide events have been investigated. Using various signal measurements as a function of range, these data have been calibrated against optical yield estimates from satellite measurements. A correction for the presence of stratospheric winds has also been applied to the observations and is found to greatly improve the relationship correlation, remaining scatter may be due to variations amongst the fireball population such as differing burst altitudes, greater or lesser amounts of fragmentation or the variability in the magnitude of stratospheric winds, which can during certain times of the year be comparable to or exceed the strength of the winds themselves. Comparison to point source, ground-level nuclear and high explosive airwave data with similar yields shows that observed bolide infrasound is consistently lower in amplitude. This downward shift relative to nuclear and HE data is interpreted as due in part to increased weak nonlinearity during signal propagation from higher altitudes. This is a likely explanation, since mean estimates of the altitude of maximum energy deposition along the bolide trajectory was found to be between 20–30 km altitude for this fireball population.

9:00

2aPA4. Numerical acoustic wave propagation in the atmosphere. Lars Ceranna (BGR, B3.11, Hannover, Germany) and Alexis Le Pichon (CEA/DASE, Bruyeres-le-Chatel, France)

A method is presented which simulates the acoustic wave-field in the atmosphere using a Chebyshev pseudo spectral approach. The aero-acoustic equation of motion is solved in spherical coordinates. The computational time is significantly reduced by applying a rotationally symmetric approximation. Since the complete wave-field is calculated for the two-dimensional model the solution provides pictures of the pressure propagation from the source to the receiver. These images make it possible to study two effects in detail: How the sound and wind profiles affect the synthetic barograms for a fixed source, and in which way the source in space and time is reflected in the synthetic recordings for a stationary elastic model. Examples for both applications are shown where measured data are compared with synthetics. The recording at the German IMS station I26DE of the September 21, 2001, chemical explosion in Toulouse is considered as a benchmark for synthetic barograms calculated for two-dimensional profiles based on NRL-G2S and MSISE/HWM. In the second example the estimation of the source functions of super-sonic moving bodies is demonstrated. The synthetic waveforms are compared with the signals of two bolides recorded at the French IMS station I24FR at Tahiti from the December 1, 2003.

9:20

2aPA5. Contribution of infrasound monitoring for atmospheric investigations. Alexis Le Pichon (CEA/DASE, BP12, Bruyeres-le-Chatel 91680, France) and Doug Drob (Naval Res. Lab., Washington, DC 20375)

The French NDC currently receives data from a number of IMS-type infrasound stations. Continuous automatic processing of the data is being performed in the [0.02–4] Hz frequency band in order to detect and characterize coherent infrasonic waves. Known and quasi permanent infrasonic sources are needed to evaluate and improve upper-wind models. Microbaroms are detected throughout the year on a global scale. Their monitoring reveals clear periodic trends in the detected bearings, providing further confirmation that long-range propagation strongly depends on the atmospheric conditions, primarily on the seasonal variability of the zonal winds. Infrasounds generated by active volcanoes also offer a unique opportunity for atmospheric studies. Recent observations in Vanuatu (New-Hebrides Islands) covering one full year show a maximal peak-to-peak azimuth deviation of ~15. Infrasound observations are used as input of an inversion procedure to evaluate more precisely the vertical structure of the wind above the stratosphere in a range of altitude inaccessible to ground based or satellite measurements. With the increasing number of IMS stations being deployed, continuing systematic investigations into infrasonic signals from volcanoes will certainly help to advance the development of operational infrasound monitoring, and significant contributions to the understanding of atmospheric propagation could be expected.

9:40

2aPA6. Detecting and characterizing infrasound signals with optical fiber infrasound sensors. Kristoffer Walker, Mark Zumberge, Jonathan Berger, Michael Hedlin (Inst. of Geophys. and Planetary Phys., Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0225), and Stephen Arrowsmith (Univ. of California, San Diego, La Jolla, CA 92093-0225)

Optical Fiber Infrasound Sensors (OFIS) are long compliant tubes wrapped with two optical fibers that interferometrically measure the differential pressure variation along the length of the tube. Because each sensor averages spatially along the length of the tube, the frequency response of the recorded pressure variation is a function of the orientation of the OFIS sensor relative to the back azimuth and incidence angle of the incoming wave. We have exploited this property to investigate the ability of various OFIS geometries to determine the back azimuth of infrasound signals. We have found that an OFIS comprised of two orthogonal 89-m-long arms having their centers separated by 63 m can resolve the back azimuth of most infrasound signals with a good signal-to-noise ratio. We find a good match between the back azimuths determined with our technique and those determined for the same signals recorded on the co-located pipe array I57US with the Progressive Multichannel Cross-Correlation technique. Based on these results and additional synthetic tests, we have built and are testing a larger OFIS with several arms that will be able to resolve signals from all directions and with small signal-to-noise ratios.

Contributed Papers

10:30

2aPA7. Listening to shower meteors with infrasound. Wayne N. Edwards (Dept. of Earth Sci., Univ. of Western Ontario, London, ON, Canada N6A 5B7), Peter G. Brown (Univ. of Western Ontario, London, ON, Canada N6A 3K7), Pavel Spurny (Astronomical Inst. of the Acad. of Sci. of the Czech Republic, Czech Republic), and Douglas O. ReVelle (Los Alamos Natl. Lab., Los Alamos, NM 87545)

Observations of the Leonid and Perseid meteor showers over the last few years in Europe using multiple camera systems, including the European Fireball Camera Network, have shown several bright meteors (abs. magnitudes -8 to -16) passing nearby to the IMS infrasound array in Freyung, Germany (I26DE). Subsequent checks of I26DE data show that these meteors were also detected infrasonically. This combination of optical location of the meteor in flight and microbarometer array beamforming has provided an excellent opportunity to delimit the altitudes at which these infrasound signals are being generated. UKMO temperature and wind data from the UARS satellite has been combined with MSIS and HWM models to reconstruct the acoustic velocity conditions present during observations. Using the reconstructed conditions, geometric ray tracing indicates that the source altitudes lie between approximately 80 to 105 kilometers; array back-azimuths appear to confirm this conclusion. This is extraordinary since meteoroids of this size, at these altitudes in the atmosphere, are in the transitional region from free-molecular flow to continuum flow. With these observations there is now confirmed infrasound from three separate meteor showers; the Leonids, Geminids and Perseids, meaning that meteor shower infrasound is much more common than previously thought.

10:45

2aPA8. Infrasound remote sensing of the upper atmosphere. Douglas Drob, Micheal Picone (E.O. Hulburt Ctr. for Space Res., Naval Res. Lab., 4555 Overlook Ave., Washington, DC 20375), and Robert Meier (School of Computational Sci., Fairfax, VA 22030)

We investigate the prospects for utilizing infrasound signals from geophysical sources to improve our knowledge of upper atmospheric winds and temperatures. We present results from numerical experiments designed to quantify the statistical performance of an idealized infrasonic upper atmospheric remote sensing network over a range of propagation modeling assumptions, source types, and geophysical variables. For a series of hypothetical events, infrasonic observables are calculated with a spherical 3-D ray tracer and series of atmospheric profiles spanning the range of geophysical parameters. These synthetic measurements are then inverted to estimate the original atmospheric background fields. A nonlinear least-squares estimation procedure that accounts for uncertainties in both the dependent and independent variables is used. The retrieved atmospheric profiles are represented by a set of truncated empirical orthogonal functions to incorporate *a priori* knowledge of atmospheric structure and reduce the number of parameters. Our numerical experiments indicate that it is indeed possible to make meaningful estimates of upper atmospheric environmental profiles with today's infrasound networks. [Work supported by the Office Naval Research.]

11:00

2aPA9. Scattering of sound and infrasound waves by internal gravity waves in the atmosphere. Vladimir E. Ostashev (NOAA/ETL 325 Broadway, Boulder, CO 80305 and Phys. Dept., New Mexico State Univ., Las Cruces, NM 88003), Igor P. Chunchuzov (Obukhov Inst. of Atmospheric Phys., Moscow, Russia), and D. Keith Wilson (U.S. Army Engineer Res. and Development Ctr., Hanover, NH 03755)

Internal gravity waves (IGW) propagating in the atmosphere produce highly anisotropic fluctuations in temperature and wind velocity that are stretched in a horizontal direction. These fluctuations can significantly affect sound propagation in a night-time boundary layer and infrasound

propagation in the stratosphere. In this paper, 3-D spectra of temperature and wind velocity fluctuations due to IGW, recently derived in the literature for the limiting case of large wave numbers, are generalized to account for small wave numbers. The generalized 3-D spectra are then used for development of a theory of sound propagation through and scattering by IGW. First, the sound scattering cross-section due to IGW is calculated. It is shown that the dependence of this scattering cross-section on parameters of the problem (the scattering angle, sound frequency, etc.) is qualitatively different from that for the case of sound scattering by isotropic turbulence. Second, the mean sound field and the transverse coherence function of a plane sound wave propagating in the atmosphere with IGW are calculated and compared with those for sound propagating through isotropic turbulence. [Work supported by ARO, Grant DAAD19-01-1-0640, and RFBR, Grant 03-05-04001.]

11:15

2aPA10. Parabolic equation (PE) model approximations and implications for infrasound. David E. Norris (BBN Technologies, 1300 N. 17th St., Ste. 400, Arlington, VA 22209, dnorris@bbn.com)

The continuous-wave parabolic equation (PE) model is widely used in the prediction of atmospheric propagation. In this study, the effects of several PE approximations are evaluated in the context of long-range infrasonic propagation. Specifically, the focus is on quantifying: phase errors resulting from different split-step Fourier (SSF) implementations, solution stability with respect to step size, and prediction sensitivity to the choice of reference sound speed. The tradeoff between improved performance gain and increased computational loading will also be considered. The study will include comparison of PE waveform predictions with measurements from infrasonic events. These comparisons are of interest in assessing the PE modeling performance, applicability, and limitations. Waveform predictions are made by integrating the continuous-wave PE model into a Fourier-synthesis Time-domain PE (TDPE).

11:30

2aPA11. Modeling volcanic infrasound propagation using the parabolic equation. Geoffrey F. Edelmann, Douglas P. Drob, Joseph F. Lingeitch, Michael D. Collins (Naval Res. Lab., 4555 Overlook Ave. SW, Washington DC, 20375), and Alexis Le Pichon (CEA/DASE BP12, 91680 Bruyeres-le Chatel, France)

Infrasound originating from volcanic sources in the 1–4 Hz range is detectable at long ranges by remote monitoring stations. The effects of wind, in addition to the atmospheric sound speed, is non-negligible for the propagation of these signals. Recent developments in parabolic equation (PE) methods are applicable to infrasonic propagation where both wide-angles and high-Mach number atmospheric flows are important. We apply this model to infrasonic data measured from volcanoes in the Archipelago of Vanuatu in Oceania (e.g., LePichon *et al.*, 2005), as well as other potentially active volcanoes. The diurnal variations and seasonal variations in measured pressure amplitudes are modeled in the context of recent multi-year observations of volcanic infrasound. The significance of upper atmospheric tidal amplitudes, range dependence, and prevailing gravity waves structure in these calculations are also considered. [Work supported by ONR.]

11:45

2aPA12. Use of improved propagation models to investigate novel infrasound signals. Joydeep Bhattacharyya and Robert Gibson (BBN Technologies, 1300 N. 17th St., #400, Arlington, VA 22209, joydeep@bbn.com)

The recent availability of improved infrasound propagation modeling tools and the installation of infrasound arrays in regions of sparse cover-

age provide new opportunities for analysis of rare signals. Analyses are presented resulting from recent natural and manmade events, and insights are discussed. Events of interest include the train explosion in Iran of February 18, 2004 and the South Asian tsunami event of December 26, 2004. Using near-real-time environmental profiles and infrasound propa-

gation models, it is shown that the predicted travel times, azimuth deviation and signal duration closely match the observed values. As the dataset of such events is small, predictive modeling can be a valuable tool in developing detection strategies for such events and may ultimately lead to a decreased incidence of false alarms in global monitoring.

TUESDAY MORNING, 17 MAY 2005

REGENCY C, 9:00 TO 11:50 A.M.

Session 2aPP

Psychological and Physiological Acoustics and Speech Communication: Current Issues in Auditory Spectral Integration

Lawrence L. Feth, Cochair

Ohio State Univ., Speech and Hearing Science, 1070 Carmack, Columbus, OH 43210-1372

Robert A. Fox, Cochair

Ohio State Univ., Speech and Hearing Science, 1070 Carmack, Columbus, OH 43210-1372

Chair's Introduction—9:00

Invited Papers

9:05

2aPP1. Spectral integration in speech and non-speech sounds. Ewa Jacewicz (Dept. Speech and Hearing Sci., Ohio State Univ., Columbus, OH 43210-1002)

Spectral integration (or formant averaging) was proposed in vowel perception research to account for the observation that a reduction of the intensity of one of two closely spaced formants (as in /u/) produced a predictable shift in vowel quality [Delattre *et al.*, *Word* **8**, 195–210 (1952)]. A related observation was reported in psychoacoustics, indicating that when the components of a two-tone periodic complex differ in amplitude and frequency, its perceived pitch is shifted toward that of the more intense tone [Helmholtz, *App. XIV* (1875/1948)]. Subsequent research in both fields focused on the frequency interval that separates these two spectral components, in an attempt to determine the size of the bandwidth for spectral integration to occur. This talk will review the accumulated evidence for and against spectral integration within the hypothesized limit of 3.5 Bark for static and dynamic signals in speech perception and psychoacoustics. Based on similarities in the processing of speech and non-speech sounds, it is suggested that spectral integration may reflect a general property of the auditory system. A larger frequency bandwidth, possibly close to 3.5 Bark, may be utilized in integrating acoustic information, including speech, complex signals, or sound quality of a violin.

9:25

2aPP2. The dispersion-focalization theory of sound systems. Jean-Luc Schwartz, Christian Abry, Louis-Jean Boë, Nathalie Vallée (Institut de la Commun. Parlée, UMR 5009 CNRS, INPG, Université Stendhal, 46 Av. Félix Viallet, 38031 Grenoble Cedex 1, France), and Lucie Ménard (Univ. du Québec à Montréal, Montréal, Canada H3C 3P8)

The Dispersion-Focalization Theory states that sound systems in human languages are shaped by two major perceptual constraints: dispersion driving auditory contrast towards maximal or sufficient values [B. Lindblom, *J. Phonetics* **18**, 135–152 (1990)] and focalization driving auditory spectra towards patterns with close neighboring formants. Dispersion is computed from the sum of the inverse squared inter-spectra distances in the (F_1 , F_2 , F_3 , F_4) space, using a non-linear process based on the 3.5 Bark critical distance to estimate F_2' . Focalization is based on the idea that close neighboring formants produce vowel spectra with marked peaks, easier to process and memorize in the auditory system. Evidence for increased stability of focal vowels in short-term memory was provided in a discrimination experiment on adult French subjects [J. L. Schwartz and P. Escudier, *Speech Comm.* **8**, 235–259 (1989)]. A reanalysis of infant discrimination data shows that focalization could well be the responsible for recurrent discrimination asymmetries [J. L. Schwartz *et al.*, *Speech Comm.* (in press)]. Recent data about children vowel production indicate that focalization seems to be part of the perceptual templates driving speech development. The Dispersion-Focalization Theory produces valid predictions for both vowel and consonant systems, in relation with available databases of human languages inventories.

9:45

2aPP3. Exploring the role of spectral integration in the perception of vowel sounds. Randy L. Diehl (Dept. of Psych. and Ctr. for Perceptual Systems, Univ. of Texas, Austin, TX 78712)

Building on the important work of Chistovich and Lublinskaya [Hearing Res. **1**, 185–195 (1979)], Syrdal [Speech Commun. **4**, 121–135 (1985)] proposed that vowel category boundaries tend to occur in quantal regions corresponding to a 3.5 Bark bandwidth of spectral integration. This proposal was shown to be consistent with acoustic distributions of certain American English vowels. Some initial perceptual results [K. A. Hoemeke and R. L. Diehl, *J. Acoust. Soc. Am.* **96**, 661–674 (1994)] also tended to support the role

of a 3.5 Bark integrator in determining major perceptual boundaries among vowel categories. However, other evidence, reviewed here, raises doubts about the generality of Syrdal's proposal. The negative evidence includes both within- and cross-language variation in the location of vowel category boundaries assumed to be governed by the 3.5 Bark integrator and the absence of elevated discriminability in putative quantal regions corresponding to the spectral limit of integration. [Work supported by NIDCD.]

10:05–10:20 Break

10:20

2aPP4. Similarities and differences in the perception of formant frequency transitions and center-of-gravity glides. Valentina Lublinskaja (Pavlov Inst. of Physiol., Russian Acad. of Sci., Makarova nab. 6, St.-Petersburg, 199034, Russia, valub@infran.ru)

In Lublinskaja (1996) the spectral “center-of-gravity” (COG) effect was observed in the identification of synthetic vowels in which the amplitude ratio of F_2 and F_3 was gradually changed in time, formant frequencies being constant. The stimuli with such virtual formant transitions (VT) were perceived as diphthong-like Russian vowels when $F_3 - F_2 < 4.5$ Bark. Within this $F_3 - F_2$ frequency separation, the stimuli with the dynamic perceptual formant (F_2^*) were identified as their equivalent vowels with real F_2 and F_3 transitions (FT). But the perception of both types of transitions was not identical. Generally, the VT signals were less salient than FT. Within smaller $F_2 - F_3$ separations ($F_2^* = 1.68$ Bark), the VT signals were first identified with 100%-accuracy as diphthong-like vowels and with larger F_2^* the glide transitions were gradually less perceptible. With separations $F_3 - F_2 > 4.5$ Bark the VT stimuli were identified as stationary vowels, and listeners' decisions were based on terminal values of F_2^* . Reduced salience of VT glides relative to FM tones was also reported for non-speech signals [Dawson and Feth, ARLO 5, 95–99 (2004)]. The talk will discuss possible mechanisms for why VT signals are less perceptually prominent than FT sounds.

10:40

2aPP5. Spectral integration in the processing of dynamic place cues for stops. Robert Allen Fox and Marc Smith (Dept. Speech and Hearing Sci., Ohio State Univ., Columbus, OH 43210-1002)

Chistovich and her colleagues have proposed the spectral “center-of-gravity” (COG) hypothesis suggesting that for static signals (such as steady-state vowels) spectral integration of two or more formant peaks may occur over a range of 3.5 Bark. More recently, studies have examined dynamic COG effects by modifying the amplitude of formants in time in order to create diphthong percepts. This presentation will review data demonstrating the COG effect in the processing of dynamic cues to place of articulation in stops in both voiced [da]-[ga] and voiceless [ta]-[ka] series. By changing only the amplitudes over time of a pair of sinewaves (in [da]-[ga]) or a pair of narrow band noise signals (in [ta]-[ka]) a “virtual F_3 glide” can be created which can serve to signal the alveolar-velar place distinction. If these signals are played to listeners without being inserted into a base token, they are heard simply as frequency glides and not as two separate tones. Listener identification and discrimination responses to CV tokens with virtual and real F_3 frequency transitions show similar overall patterns, but virtual transitions are less salient place cues than real transitions. An auditory model incorporating the dynamic COG effect will be described.

11:00

2aPP6. Auditory spectral integration of dynamic signals. Lawrence Feth (Dept. of Speech and Hearing Sci., Ohio State Univ., Columbus, OH 43210, feth.1@osu.edu), Ashok Krishnamurthy, and Daniel Hack (Ohio State Univ., Columbus, OH 43210)

Auditory spectral integration often refers to the ability of the auditory system to sum information across a wide frequency range to improve the detectability of a supra-critical bandwidth signal in a masking experiment. However, in other paradigms, the spectral components of these broad bandwidth signals can be resolved by auditory system. These findings lead to a resolution versus integration paradox for the frequency domain similar to the resolution versus integration paradox in the time domain that led to the multiple-looks hypothesis [N. F. Viemeister and G. H. Wakefield, J. Acoust. Soc. Am., 90, 858–865 (1991)]. In speech, spectral integration refers to finding that two or more resonance peaks in a synthetic vowel can be approximated by a single peak located at the spectral center-of-gravity (COG). In 1996, Lublinskaja demonstrated that changing the COG of a two-resonance signal over time leads listeners to hear a virtual frequency transition that follows the dynamic COG. Dynamic signals may require the auditory system to take multiple looks in both time and frequency. This talk will present a series of psychoacoustic experiments using dynamic signals and a computational model that incorporates peripheral auditory processing and a modification of the PSC model.

Contributed Papers

11:20

2aPP7. Difference between diotic and dichotic presentations on audiovisual synchronization. Rie Nakamura and Kiyooki Aikawa (School of Media Sci., Tokyo Univ. of Technol., 1404-1 Katakuracho, Hachioji, Tokyo 192-0982, Japan, aik@media.teu.ac.jp)

Significant differences were found between diotic and dichotic presentations of auditory stimuli on audiovisual synchronization. Two types of moving sphere were used for visual stimuli; (a) a sphere moved straight from left to right at a constant velocity, and (b) it moved up and right and then turned down and right at the midpoint. Auditory stimuli included (1) a continuous pure tone (1000 Hz), (2) a sequence of two short tones of 30

ms (1000 Hz), (3) a sweep tone (500 Hz → 1500 Hz), and (4) a Frequency-Modulated tone (500 Hz → 1500 Hz → 500 Hz). The duration of these stimuli was 240 ms each. Visual stimuli were presented on a Liquid-Crystal Display. Auditory stimuli were presented by headphones. The time difference between auditory and visual stimuli was changed between –360 ms and +360 ms. The sound source linearly moved from left to right in case of dichotic presentation. Subjective tests, using the forced-choice of synchronized or not, clarified the significant differences between diotic and dichotic presentations when the auditory stimulus (1) or (4) was presented with the visual stimulus (b). No significance was obtained for other combinations of auditory and visual stimuli. [Work supported by NTT.]

2aPP8. Frequency band-importance functions for auditory and auditory-visual speech recognition. Ken W. Grant (Walter Reed Army Medical Ctr., Army Audiol. and Speech Ctr., Washington, DC 20307-5001)

In many everyday listening environments, speech communication involves the integration of both acoustic and visual speech cues. This is especially true in noisy and reverberant environments where the speech signal is highly degraded, or when the listener has a hearing impairment. Understanding the mechanisms involved in auditory-visual integration is a primary interest of this work. Of particular interest is whether listeners are able to allocate their attention to various frequency regions of the speech signal differently under auditory-visual conditions and auditory-alone con-

ditions. For auditory speech recognition, the most important frequency regions tend to be around 1500–3000 Hz, corresponding roughly to important acoustic cues for place of articulation. The purpose of this study is to determine the most important frequency region under auditory-visual speech conditions. Frequency band-importance functions for auditory and auditory-visual conditions were obtained by having subjects identify speech tokens under conditions where the speech-to-noise ratio of different parts of the speech spectrum is independently and randomly varied on every trial. Point biserial correlations were computed for each separate spectral region and the normalized correlations are interpreted as weights indicating the importance of each region. Relations among frequency-importance functions for auditory and auditory-visual conditions will be discussed.

TUESDAY MORNING, 17 MAY 2005

REGENCY D, 8:00 A.M. TO 12:00 NOON

Session 2aSC

Speech Communication: Language Acquisition, Bilingual Studies and Category Learning (Poster Session)

Murray J. Munro, Chair

Dept. of Linguistics, Simon Fraser Univ., 8888 University Dr., Burnaby, BC V5A 1S6, Canada

Contributed Papers

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon.

2aSC1. L1 literacy affects L2 pronunciation intake and text vocalization. Martin Walton (Dept. Langues, Université Montesquieu Bordeaux 4, Ave. Léon Duguit, 33608 PESSAC Cedex, France)

For both deaf and hearing learners, L1 acquisition calls on auditive, gestural and visual modes in progressive processes over longer stages imposed in strictly anatomical and social order from the earliest pre-lexical phase [Jusczyk (1993), Kuhl & Meltzoff (1996)] to ultimate literacy. By contrast, L2 learning will call on accelerating procedures but with restricted input, arbitrated by L1 literacy as can be traced in the English of French-speaking learners, whether observed in spontaneous speech or in text vocalization modes. An inventory of their predictable omissions, intrusions and substitutions at suprasegmental and syllabic levels, many of which they can actually hear while unable to vocalize in real-time, suggests that a photogenic segmentation of continuous speech into alphabetical units has eclipsed the indispensable earlier phonogenic module, filtering L2 intake and output. This competing mode analysis hypothesizes a critical effect on L2 pronunciation of L1 graphemic procedures acquired usually before puberty, informing data for any Critical Period Hypothesis or amounts of L1 activation influencing L2 accent [Flege (1997, 1998)] or any psychoacoustic French deafness with regard to English stress-timing [Dupoux (1997)]. A metaphonic model [Howell & Dean (1991)] adapted for French learners may remedially distance L1 from L2 vocalization procedures.

2aSC2. Lexical representation of novel L2 contrasts. Rachel Hayes-Harb (Dept. of Linguist., Univ. of Utah, 255 S. Central Campus Dr., Rm. 2328, Salt Lake City, UT 84112-0492, hayes-harb@linguistics.utah.edu) and Kyoko Masuda (Georgia Inst. of Technol., Atlanta, GA 30332-0375)

There is much interest among psychologists and linguists in the influence of the native language sound system on the acquisition of second languages (Best, 1995; Flege, 1995). Most studies of second language

(L2) speech focus on how learners perceive and produce L2 sounds, but we know of only two that have considered how novel sound contrasts are encoded in learners' lexical representations of L2 words (Pallier *et al.*, 2001; Ota *et al.*, 2002). In this study we investigated how native speakers of English encode Japanese consonant quantity contrasts in their developing Japanese lexicons at different stages of acquisition (Japanese contrasts singleton versus geminate consonants but English does not). Monolingual English speakers, native English speakers learning Japanese for one year, and native speakers of Japanese were taught a set of Japanese nonwords containing singleton and geminate consonants. Subjects then performed memory tasks eliciting perception and production data to determine whether they encoded the Japanese consonant quantity contrast lexically. Overall accuracy in these tasks was a function of Japanese language experience, and acoustic analysis of the production data revealed non-native-like patterns of differentiation of singleton and geminate consonants among the L2 learners of Japanese. Implications for theories of L2 speech are discussed.

2aSC3. L2 use and stimulus complexity in perceived accent ratings. Astrid Zerla Doty, Ruth Huntley Bahr (Dept. of Commun. Sci. and Disord., Univ. of South Florida, PCD 1017, 4202 E. Fowler Ave., Tampa, FL 33620, rbahr@chuma1.cas.usf.edu), and Judith Becker Bryant (Univ. of South Florida, Tampa, FL 33620)

Listener perception of accentedness has been shown to be influenced by experience with L2 (measured by length of residence in US). However, frequency of L1 use and degree of linguistic complexity (defined by the number of non-native phonetic features targeted) may provide more insight into the role of experience in the perception of accentedness. Three groups of listeners (monolingual English and Spanish [L1] speakers divided into two groups of high and low use of English [L2]) rated the accentedness of bilingual speakers who spoke with varying degrees of accentedness. The speakers read sentences adapted from Magan (1998) to

include linguistic aspects likely to be difficult for native Spanish speakers. Listeners performed similarly in rating speakers degree of accent. Amount of daily L1 use only influenced the ratings of the slightly accented group; the high-use bilingual group rated these speakers as more accented than the native English group, regardless of level of linguistic difficulty. These results suggest that the high-use groups lack of L2 experience made them less perceptually sensitive to certain phonetic features of English. Because speakers did not make the predicted target errors, the listener groups may have based their ratings on features not targeted in this investigation.

2aSC4. Training for learning Mandarin tones: A comparison of production and perceptual training. Xinchun Wang (Dept. of Linguist., California State Univ., Fresno, 5245 N. Backer Ave., M/S PB 92, Fresno, CA 93740)

Mandarin Chinese lexical tones pose difficulties for non-native speakers whose first languages contrast or do not contrast lexical tones. In this study, both tone language and non-tone language speaking learners of Mandarin Chinese were trained for three weeks to identify the four Mandarin lexical tones. One group took the production training with both visual and audio feedback using Kay Sona Speech II software. The target tones produced by native Mandarin speakers were played back through a pair of headphones and the pitch contours of the target tones were displayed on the computer screen on the top window to be compared with the trainees productions which appear in real time in the bottom window. Another group of participants took the perceptual training only with four-way forced choice identification tasks with immediate feedback. The same training tokens were used in both training modes. Pretest and post test data in perception and production were collected from both groups and were compared for effectiveness of training procedures.

2aSC5. Production quality of /r/ and /l/ liquids among Cantonese and Mandarin ESL learners. Donald Derrick (Dept. of Linguist., Univ. of British Columbia, 1866 Main Mall, Buchanan E270, UBC, Vancouver, BC, Canada V6T 1Z1, dderrick@interchange.ubc.ca)

Perceptual interference theories suggest L2 language learners produce phonemes based on their native language phonology (Iverson *et al.*, 2004, Cognition). This present study investigated the impact of differing native language segmental inventories on the acquisition of the English r/l contrast. Northern Mandarin dialects exhibit coda /r/ similar to the English bunched /r/, while Cantonese exhibits no r-like liquids (Gick *et al.*, 2003, under review). The Mandarin segmental inventory provides more of a basis for acquiring the English r/l contrast than the Cantonese inventory. It is therefore predicted that Mandarin speakers will acquire the r/l contrast with a lower level of experience with English than the Cantonese speakers. One Cantonese and two Mandarin ESL learners produced r/l sounds in minimally contrastive English words in simple and complex onset, coda, and intervocalic positions. The data were analyzed in two ways. Four native English listeners were asked to judge for each word whether the target consonant was /r/ or /l/. Also, ultrasound tongue images were analyzed for component /r/ and /l/ gestures. Results provided partial support for the hypothesis. Implications for theories of second language acquisition will be discussed.

2aSC6. Native and non-native perception of phonemic length contrasts in Japanese: Effects of speaking rate and presentation context. Amanda Wilson, Hiroaki Kato (ATR Human Information Sci. Labs., Kyoto 619-0288, Japan, kato@atr.jp), and Keiichi Tajima (Hosei Univ., Tokyo 102-8160, Japan)

Japanese words can be distinguished by the length of phonemes, e.g., “chizu” (map) versus “chiizu” (cheese). Perceiving these length contrasts is therefore important for learning Japanese as a second language. The present study examined native English listeners’ perception of length contrasts at different speaking rates and in different contexts. Stimuli con-

sisted of 20 Japanese word pairs that minimally contrasted in vowel length, and 10 synthesized nonwords. The nonwords were created by modifying the duration of the second vowel of the nonword “erete” along a continuum (from “erete” to “ereete”). Stimuli were presented with or without a carrier sentence at three rates (fast, normal, slow). Rate was either fixed or randomized trial by trial. Sixteen native English and 16 native Japanese listeners participated in a single-stimulus, two-alternative forced-choice identification task. Results suggest that native Japanese listeners’ identification boundaries systematically shifted due to changes in speaking rate when the stimuli were in the context of a sentence with mixed rates of presentation. In contrast, native English listeners show a shift in the opposite direction, suggesting that they did not follow the variation in speaking rate. These results will be discussed from the viewpoint of training second-language phoneme perception. [Work supported by JSPS.]

2aSC7. Children’s abilities to distinguish novel languages. Z. S. Bond and Verna Stockmal (Dept. of Linguist., Ohio Univ., Athens, OH 45701)

When adults hear spoken samples of a language which they do not know, they can often identify it and discriminate between languages even when produced by the same talkers. Children have much less experience making metalinguistic judgments. How do children respond to languages which they do not know? We have conducted three experiments examining the abilities of 4-year old and 8-year old children to discriminate between spoken samples of different languages produced by bilingual talkers. We constructed listening tests from 5-second phrases excerpted from fluent reading provided by the talkers. In the three experiments, we progressively simplified the response mode employed by the children as well as the cognitive load of the task. Even in the simplest version, only a third of the 4-year-olds could do the task while the 8-year old children performed above chance in all three experiments. The younger children tended to respond different more than same, as if their criterion for same was identify.

2aSC8. Innovative /ye/ and /we/ sequences in recent loans in Japanese. Timothy Vance and Yuka Matsugu (Univ. of Arizona, East Asian Studies, P.O. Box 210105, Tucson, AZ 85721-0105)

The GV sequences /ye/ and /we/ do not occur in Japanese except perhaps in recent loans. Katakana spellings of the relevant loans in authoritative dictionaries are inconsistent, and it is not clear whether native speakers treat them as containing the GV sequences /ye/ and /we/ or as containing the VV sequences /ie/ and /ue/. Native speakers of Japanese with minimal exposure to spoken English were recorded producing some relevant loans in response to picture prompts. The same speakers were also recorded producing some native words containing uncontroversial /ie/ and /ue/ sequences. All the productions are being analyzed acoustically to determine whether they show the expected contrast between GV and VV sequences. A VV sequence is disyllabic (and bimoraic) and should therefore have greater duration and more gradual formant movements than a monosyllabic (and monomoraic) GV sequence. Utterance-initially, a VV sequence should have a LH pitch pattern and should be preceded by a nondistinctive glottal stop, whereas a GV sequence should have a H pitch pattern and should have smooth onset.

2aSC9. Categorization of speech sounds by Norwegian/English bilinguals. Audny T. Dypvik and Elzbieta B. Slawinski (Psych. Dept., Univ. of Calgary, 2500 University Dr., Calgary, AB, Canada T2N 1N4, atdypvik@ucalgary.ca)

Bilinguals who learned English late in life (late bilinguals) as opposed to those who learned English early in life (early bilinguals) differ in their perception of phonemic distinctions. Age of acquisition of a second language as well as depth of immersion into English is influenced by perceptual differences of phonemic contrasts between monolinguals and bilin-

guals, with consequences for speech production. The phonemes /v/ and /w/ are from the same category in Norwegian, rendering them perceptually indistinguishable to the native Norwegian listener. In English, /v/ and /w/ occupy two categories. Psychoacoustic testing on this phonemic distinction in the current study will compare perceptual abilities of monolingual English and bilingual Norwegian/English listeners. Preliminary data indicates that Norwegian/English bilinguals demonstrate varying perceptual abilities for this phonemic distinction. A series of speech sounds have been generated by an articulatory synthesizer, the Tube Resonance Model, along a continuum between the postures of /v/ and /w/. They will be presented binaurally over headphones in an anechoic chamber at a sound pressure level of 75 dB. Differences in the perception of the categorical boundary between /v/ and /w/ among English monolinguals and Norwegian/English bilinguals will be further delineated.

2aSC10. Perception of coarticulated tones by non-native listeners.

Tessa Bent (2016 Sheridan Rd., Evanston, IL 60208, t-bent@northwestern.edu)

Mandarin lexical tones vary in their acoustic realization depending on the surrounding context. Native listeners compensate for this tonal coarticulation when identifying tones in context. This study investigated how native English listeners handle tonal coarticulation by testing native English and Mandarin listeners discrimination of the four Mandarin lexical tones in tri-syllabic sequences in which the middle tone varied while the first and last tones were held constant. Three different such frames were tested. As expected, Mandarin listeners discriminated all pairs in all contexts with a high degree of accuracy. English listeners exhibited poorer discrimination than Mandarin listeners and their discrimination accuracy showed a high degree of context dependency. In addition to assessing accuracy, reaction times to correctly discriminated different trials were entered into a multidimensional scaling analysis. For both listener groups, the arrangement of tones in perceptual space varied depending on the surrounding context suggesting that listeners attend to different acoustic attributes of the target tone depending on the surrounding tones. These results demonstrate the importance for models of cross-language speech perception of including contextual variation when characterizing the perception of non-native prosodic categories. [Work supported by NIH/NIDCD]

2aSC11. Plasticity in speech production and perception: A study of accent change in young adults.

Bronwen G. Evans and Paul Iverson (Dept. of Phonet. and Linguist., Univ. College London, Wolfson House, 4, Stephenson Way, London, NW1 2HE, UK bron@phon.ucl.ac.uk)

This study investigated plasticity in speech production and perception among university students, as individuals change their accent from regional to educated norms. Subjects were tested before beginning university, 3 months later and on completion of their first year of study. At each stage they were recorded reading a set of test words and a short passage. They also completed two perceptual tasks; they found best exemplar locations for vowels embedded in carrier sentences and identified words in noise. The results demonstrated that subjects changed their spoken accent after attending university. The changes were linked to sociolinguistic factors; subjects who were highly motivated to fit in with their university community changed their accent more. There was some evidence for a link between production and perception; between-subject differences in production and perception were correlated. However, this relationship was weaker for within-subject changes in accent over time. The results suggest that there were limitations in the ability of these subjects to acquire new phonological rules.

2aSC12. Articulatory settings of French-English bilingual speakers.

Ian Wilson (Dept. of Linguist., Univ. of British Columbia, 1866 Main Mall, Buchanan E270, Vancouver, BC, Canada V6T 1Z1, ilwilson@interchange.ubc.ca)

The idea of a language-specific articulatory setting (AS), an underlying posture of the articulators during speech, has existed for centuries [Laver, *Historiogr. Ling.* 5 (1978)], but until recently it had eluded direct measurement. In an analysis of x-ray movies of French and English monolingual speakers, Gick *et al.* [Phonetica (in press)] link AS to inter-speech posture, allowing measurement of AS without interference from segmental targets during speech, and they give quantitative evidence showing AS to be language-specific. In the present study, ultrasound and Optotrak are used to investigate whether bilingual English-French speakers have two ASs, and whether this varies depending on the mode (monolingual or bilingual) these speakers are in. Specifically, for inter-speech posture of the lips, lip aperture and protrusion are measured using Optotrak. For inter-speech posture of the tongue, tongue root retraction, tongue body and tongue tip height are measured using optically-corrected ultrasound. Segmental context is balanced across the two languages ensuring that the sets of sounds before and after an inter-speech posture are consistent across languages. By testing bilingual speakers, vocal tract morphology across languages is controlled for. Results have implications for L2 acquisition, specifically the teaching and acquisition of pronunciation.

2aSC13. Paired variability indices in assessing speech rhythm in Spanish/English bilingual language acquisition.

Richard Work, Jean Andruski, Eugenia Casielles, Sahyang Kim (Wayne State Univ., Detroit, MI 48202), and Geoff Nathan (Wayne State Univ., Detroit, MI 48202)

Traditionally, English is classified as a stress-timed language while Spanish is classified as syllable-timed. Examining the contrasting development of rhythmic patterns in bilingual first language acquisition should provide information on how this differentiation takes place. As part of a longitudinal study, speech samples were taken of a Spanish/English bilingual child of Argentinean parents living in the Midwestern United States between the ages of 1;8 and 3;2. Spanish is spoken at home and English input comes primarily from an English day care the child attends 5 days a week. The parents act as interlocutors for Spanish recordings with a native speaker interacting with the child for the English recordings. Following the work of Grabe, Post and Watson (1999) and Grabe and Low (2002) a normalized Pairwise Variability Index (PVI) is used which compares, in utterances of minimally four syllables, the durations of vocalic intervals in successive syllables. Comparisons are then made between the rhythmic patterns of the child's productions within each language over time and between languages at comparable MLUs. Comparisons are also made with the rhythmic patterns of the adult productions of each language. Results will be analyzed for signs of native speaker-like rhythmic production in the child.

2aSC14. Recognition of function words in 8-month-old French-learning infants.

Rushen Shi and Bruno Gauthier (Dept. of Psych., Univ. of Quebec in Montreal, C.P. 8888, Succ. Ctr.-Ville, Montreal, QC, Canada H3C 3P8)

Previous work has shown that German-learning 7-9-month-old infants recognize function words (Hoehle and Weissenborn, 2003). English-learning infants recognize function words around 10.5-11 months (Schafer *et al.* 1998; Shady, 1996; Shi *et al.*, 2003, 2004), and the highly frequent determiner "the" at 8 months (Shi *et al.*, 2004). The present study investigates French-learning infants' recognition of function words. As French is a syllable-timing language, the fuller syllabic status may allow infants to recognize function words earlier than English-learning infants. Syntactically and morphologically, functional elements occur more systematically in French than in English, providing reliable statistical cues to functor segmentation. Using a preferential looking procedure, we familiarized 8-month-olds with a target function word ("des," "la," "mes" or "ta"), and tested them with phrases containing the target versus a non-target.

Results showed that infants' looking time to the phrases containing the targets versus those containing the non-targets differed significantly. Thus, infants recognized the target functors in continuous speech. As the targets included both high-frequency ("des," "la") and low-frequency ("mes," "ta") function words, we suggest that infants may begin segmenting high-frequency functors at an even younger age. The implications of early processing of function words to language acquisition will be discussed.

2aSC15. Bilinguals reaction times and category goodness judgments in two language-sets: Spanish and English. Adrian Garcia-Sierra (Univ. of Texas, CMA A2-2001, Univ. Station A1100, Austin, TX 78712, gasa@austin.utexas.edu)

Two monolingual groups (Spanish and English speakers) and one bilingual group (Mexican-American) were compared in reaction times (RTs) and category goodness (CG) ratings during the identification of a synthetic continuum ranging from /ga/ to /ka/. The task consisted of two sessions for the bilingual listeners and one session for the monolingual listeners. Each session was performed in either a Spanish or English language-set. Language-sets were defined by the language used by the researcher and by filler sentences in the language of interest during the identification task (e.g., what do you hear? or Qu fue lo que escuchó?). The results show that voicing boundaries were different across all groups including bilinguals in both language-sets (voicing boundary shift). RTs were larger near the voicing boundaries in all groups. Interestingly, bilinguals showed two RT peaks, whereas monolinguals showed only one RT peak. In regard to CG, participants were asked to rate how good exemplar each token was (1 = worst; 5 = best). Bilinguals rated the synthetic continuum in both language-sets. The scores given by the bilinguals differed in each language-set at VOT values close to the voicing boundary.

2aSC16. Perceived nativeness and sensitivity to temporal adjustments in speech. Yue Wang (Dept of Linguist., Simon Fraser Univ., Burnaby, BC, Canada V5A 1S6, yuew@sfu.ca) and Dawn M. Behne (Norwegian Univ. of Sci. and Technol., NO-7491 Trondheim, Norway)

Native Mandarin Chinese speakers productions of English consonant-vowel (CV) syllables have shown syllable-internal temporal adjustments in the direction of native (English)-like CVs (Wang and Behne, 2004). The current study presents two experiments investigating whether these temporal adjustments affect perceived nativeness. For three production types (native-English, Chinese productions of English, native-Chinese), three syllable-internal timing patterns (English-like, Chinese-English-like, Chinese-like) were applied, resulting in nine stimuli types. Native English listeners judged how English-like each stimulus was on a 7-point scale. In the first experiment, production-types and timing patterns were randomized. Results show that listeners can reliably identify nativeness of the three productions, with Chinese productions of English perceived as intermediate to the native Chinese and native American English productions. Listeners also showed a tendency toward using timing within the CV to identify nativeness. In the second experiment the same materials were therefore blocked by production type. Results reveal the perceptual saliency of the temporal adjustments in nonnative productions. These findings support a view of L2 acquisition as a gradual process toward the target L2 (e.g., Caramazza *et al.*, 1973). The current study extends this view, showing evidence that listeners can perceive the inter-language system, bearing the nature of both L1 and L2.

2aSC17. Relation between perception and production ability during a speech training course. Teruaki Tsumahima and Mayumi Hamada (Ryutsu Kagaku Univ., 3-1, Gakuen-Nishi-Machi, Nishi-Ku, Kobe, 651-2188, Japan)

Previous research has indicated that perception and production ability are significantly related among L2 learners in natural learning settings. The present study focused on the relation between these two abilities

during speech training, specifically examining whether perception or production training alone led to improvements of production or perception ability, and whether the order of perception and production training influenced the improvement of the abilities. Forty-five Japanese university students received thirteen-week perception and production training on six American English contrasts (e.g., b/v, l/r). One group received perception training first, followed by production training, while the order of training was reversed for the other group. It was found that the initial perception training in the former group resulted in a significant but relatively weak transfer to improvements of production ability, while the initial production training in the latter group showed a significant and strong transfer to those of perception ability. It was also shown that the improvements of perception and production ability during the speech training were significantly related in both groups. The order of training, however, did not significantly influence the overall degree of improvements in either of the abilities. Implications for speech learning models will be discussed.

2aSC18. The effect of L1 prosodic backgrounds of Cantonese and Japanese speakers on the perception of Mandarin tones after training. Connie K. So (Dept. of Linguist., Simon Fraser Univ., Burnaby, BC, Canada V5A 1S6, kso@sfu.ca)

The present study investigated to what extent ones' L1 prosodic backgrounds affect their learning of a new tonal system. The question as to whether native speakers of a tone language perform differently from those of a pitch accent language will be addressed. Twenty native speakers of Hong Kong Cantonese (a tone language) and Japanese (a pitch accent language) were assigned to two groups. All of them had had no prior knowledge of Mandarin, and had never received any form of musical training before they participated in the study. Their performance of the identification of Mandarin tones before and after a short-term training was compared. Analysis of listeners' tonal confusions in the pretest, posttest, and generalization tests revealed that both Cantonese and Japanese listeners had more confusion for two contrastive tone pairs: Tone 1–Tone 4, and Tone 2–Tone 3. Moreover, Cantonese speakers consistently had greater difficulty than Japanese speakers in distinguishing the tones in each pair. These imply that listeners L1 prosodic backgrounds are at work during the process of learning a new tonal system. The findings will be further discussed in terms of the Perceptual Assimilation Model (Best, 1995). [Work supported by SSHRC.]

2aSC19. How auditory discontinuities and linguistic experience affect the perception of speech and non-speech in English- and Spanish-speaking listeners. Jessica F. Hay (Dept. Psych., Univ. of Texas, Austin, TX 78712, hay@psy.utexas.edu), Lori L. Holt (Carnegie Mellon Univ., Pittsburgh, PA 15213), Andrew J. Lotto (Boys Town Natl. Res. Hospital, Omaha, NE 68131), and Randy L. Diehl (Univ. of Texas, Austin, TX 78712)

The present study was designed to investigate the effects of long-term linguistic experience on the perception of non-speech sounds in English and Spanish speakers. Research using tone-onset-time (TOT) stimuli, a type of non-speech analogue of voice-onset-time (VOT) stimuli, has suggested that there is an underlying auditory basis for the perception of stop consonants based on a threshold for detecting onset asynchronies in the vicinity of +20 ms. For English listeners, stop consonant labeling boundaries are congruent with the positive auditory discontinuity, while Spanish speakers place their VOT labeling boundaries and discrimination peaks in the vicinity of 0 ms VOT. The present study addresses the question of whether long-term linguistic experience with different VOT categories affects the perception of non-speech stimuli that are analogous in their acoustic timing characteristics. A series of synthetic VOT stimuli and TOT stimuli were created for this study. Using language appropriate labeling and ABX discrimination tasks, labeling boundaries (VOT) and discrimination peaks (VOT and TOT) are assessed for 24 monolingual English speakers and 24 monolingual Spanish speakers. The interplay between language experience and auditory biases are discussed. [Work supported by NIDCD.]

2aSC20. Method for automatic measurement of second language speaking proficiency. Jared Bernstein and Jennifer Balogh (Ordnate Corp., Menlo Park, CA 94025)

Spoken language proficiency is intuitively related to effective and efficient communication in spoken interactions. However, it is difficult to derive a reliable estimate of spoken language proficiency by situated elicitation and evaluation of a person's communicative behavior. This paper describes the task structure and scoring logic of a group of fully automatic spoken language proficiency tests (for English, Spanish and Dutch) that are delivered via telephone or Internet. Test items are presented in spoken form and require a spoken response. Each test is automatically-scored and primarily based on short, decontextualized tasks that elicit integrated listening and speaking performances. The tests present several types of tasks to candidates, including sentence repetition, question answering, sentence construction, and story retelling. The spoken responses are scored according to the lexical content of the response and a set of acoustic base measures on segments, words and phrases, which are scaled with IRT methods or parametrically combined to optimize fit to human listener judgments. Most responses are isolated spoken phrases and sentences that are scored according to their linguistic content, their latency, and their fluency and pronunciation. The item development procedures and item norming are described.

2aSC21. Effects of linguistic experience on early levels of perceptual tone processing. Tsan Huang (Dept. of Linguist., SUNY Buffalo, 609 Baldy Hall, Buffalo, NY 14260) and Keith Johnson (UC Berkeley, Berkeley, CA 94720-2650)

This study investigated the phenomenon of language-specificity in Mandarin Chinese tone perception. The main question was whether linguistic experience affects the earliest levels of perceptual processing of tones. Chinese and American English listeners participated in four perception experiments, which involved short inter-stimulus intervals (300 ms or 100 ms) and an AX discrimination or AX degree-of-difference rating task. Three experiments used natural speech monosyllabic tone stimuli and one experiment used time-varying sinusoidal simulations of Mandarin tones. AE listeners showed psychoacoustic listening in all experiments, paying much attention to onset and offset pitch. Chinese listeners showed language-specific patterns in all experiments to various degrees, where tonal neutralization rules reduced perceptual distance between two otherwise contrastive tones for Chinese listeners. Since these experiments employed procedures hypothesized to tap the auditory trace mode (Pisoni, *Percept. Psychophys.* **13**, 253–260 (1973)), language-specificity found in this study seems to support the proposal of an auditory cortical map [Guenther *et al.*, *J. Acoust. Soc. Am.* **23**, 213–221 (1999)]. But the model needs refining to account for different degrees of language-specificity, which are better handled by Johnsons (2004, *TLS03:26-41*) lexical distance model, although the latter model is too rigid in assuming that linguistic experience does not affect low-level perceptual tasks such as AX discrimination with short ISIs.

2aSC22. Perceptual and production variables in explicating interlanguage speech intelligibility benefit. Ameer P. Shah and Zoi Vavva (Dept. of Speech and Hearing, Cleveland State Univ., 2121 Euclid Ave., MC 431-B, Cleveland, OH 44115, a.shah101@csuohio.edu)

This study attempts to investigate the importance of the degree of similarity or difference in the language backgrounds of the speakers and listeners, as it interacts differentially in intelligibility judgment of foreign-accented speech (Bent and Bradlow, 2003). The present study attempts to clarify the distinction in the matched and mismatched listening conditions, in context of addressing the overarching question whether auditory exposure to a language alone, without corresponding proficiency in production of that language, can provide a listening advantage. Particularly, do listeners understand accented-English speech spoken by native individuals of the language to which they are exposed to, as compared to listeners without that exposure? Greek-accented English speakers (and native monolin-

gual English speakers) were judged for their speech intelligibility by four groups of listeners ($n=10$, each): native Greek speakers (matched), Greek-Americans (matched only through auditory exposure to Greek without any corresponding spoken proficiency), native monolingual American-English speakers (unmatched), and a mixed group (mismatched). Pilot data have shown that the intelligibility judgments by Greek-American listeners are intermediate to the native Greeks, and both the American-English and the mixed group. Further data-collection is underway, and will be presented as they bear important theoretical and clinical implications.

2aSC23. Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals. Kyoung-Ho Kang and Susan G. Guion (Dept. of Linguist., 1290 Univ. of Oregon, Eugene, OR 97403, kkang@darkwing.uoregon.edu)

The stop systems of adult Korean (L1)-English (L2) bilinguals were studied through acoustic analysis of Korean and English stop productions. The interaction of the first and second language stop systems was investigated as a function of age of exposure to English. The main goals of the investigation were to examine the extent to which early and the late bilinguals produced a given stop category in a native-like way and the extent to which the two stop systems were held independent from each other in the phonological systems of the bilinguals. Two specific questions were asked: whether early bilinguals were more native-like in the production of English stops and thus were more likely to establish L2 phonetic categories than late bilinguals, and whether the Early bilinguals maintain a greater extent of independence between the Korean and English stop systems than the late bilinguals. For this purpose, productions of Korean and English stops were analyzed in terms of three acoustic-phonetic properties, VOT, $H1-H2$, and f_0 . The results suggested that fine-grained phonetic information is crucial to investigation of bilingual phonological systems and they were discussed in terms of the role of perceived phonetic distance between L1 and L2 phonetic categories in L2 speech learning.

2aSC24. Validation of an automatic measurement of Spanish speaking proficiency. Elizabeth Rosenfeld, Jared Bernstein, and Jennifer Balogh (Ordnate Corp., Menlo Park, CA)

A 15-min computer-based test of spoken Spanish was designed to measure candidate proficiency in Spanish. The test presents seven tasks: reading, elicited imitation, word opposites, short-answer questions, sentence constructions, opinion questions, and story retellings. The tests were presented to 579 adult non-native Spanish learners and to 552 native Spanish speakers. Expert human judgments of the non-native responses showed that the spoken response material carried sufficient information for highly reliable judgments of proficiency. In the development and validation process, 57 000 responses were transcribed and 21 000 human judgments were analyzed. The paper describes the validation of the automatic scoring system with reference to concurrent oral proficiency interviews conducted by professional raters certified by the US Government or by ACTFL. The outcomes of the comparisons of the machine scored tests with interactive human interviews and with human ratings from recorded speech indicate that the test produces scores that have virtually the same information that is found in oral proficiency interviews. Almost all assessments correlate highly with the other assessments with coefficients in the range 0.86–0.96. The test correlation with the combined interview scores ($r=0.92$) is higher than the inter-rater reliability of the professional interviewers themselves.

2aSC25. An acoustic investigation of the Cantonese vowels in the speech of the adult and child speakers. Wai-Sum Lee (Dept. of Linguist., The Univ. of Hong Kong, Pokfulam Rd., Hong Kong, wsleeba@hku.hk)

The study analyzes the formant center frequencies for the seven Cantonese vowels [i, y, u, e, æ, ɔ, a] from 30 native speakers of Cantonese, 10 male and 10 female adults and 5 male and 5 female 9–10 year old children. Results show that the formant frequencies for the vowels are

largest for the female children, followed by the male children, female adults, and male adults in decreasing order. Despite the differences, the patterns of formant frequencies for any one vowel for the different groups are similar. The difference in F -values for any one vowel between the male and female children is smaller than the difference between the male and female adults. As for individual formant frequencies, the difference in $F1$ between the males and females of the same age group and between the adults and children of the same gender group is smaller for the high vowels [i, y, u] than the non-high vowels [ɛ, æ, ɔ, a]. The difference in $F2$ between the males and females of the same age group and between the adults and children of the same gender group is smaller for the high rounded vowels [y, u] than the other vowels. The paper will also present the ratios of speaker group-to-speaker group for individual formant frequencies.

2aSC26. Voice onset time (VOT) in Canadian French and English: Monolingual and bilingual adults. Andrea A. N. MacLeod and Carol Stoel-Gammon (Dept. of Speech and Hearing Sci., Univ. of Washington, 1417 N.E. 42nd St., Seattle, WA 98119)

This study focused on the contrasts produced by early bilingual speakers ($n=6$) across their two languages in comparison with monolingual speakers (Canadian English (CE), $n=5$; Canadian French (CF), $n=6$). VOT production was measured in monosyllabic CE and CF words that began with one of four stop consonants, /p, b, t, d/ followed by one of three vowels. A total of 14–18 words for each of the four stop consonants for each language was elicited with a total number 1700 acoustically analyzed productions. The participants were tested individually in quiet rooms using a single target language throughout the session. As expected, the monolingual speakers produced a two-way contrast (statistically significant: $p < 0.05$): for CE speakers, short-lag VOT versus long-lag VOT; for CF speakers, lead VOT versus short-lag VOT. Rather than producing a two-way contrast (e.g., lead VOT versus lag VOT) or a three-way contrast (e.g., lead VOT versus short-lag VOT versus long-lag VOT), the bilingual speakers produced a four-way contrast (statistically significant: $p < 0.05$): long lead VOT (CF /b, d/), short lead VOT (CE /b, d/), short-lag VOT (CF /p, t/) and long-lag VOT (CE /p, t/). These results suggest that bilinguals are maintaining phonetic contrasts both within and across their two languages.

2aSC27. The effect of speaking rate on perception of syllables in second-language speech. Keiichi Tajima (Dept. of Psych., Hosei Univ., Tokyo 102-8160, Japan; ATR Human Information Sci. Labs., Kyoto 619-0288, Japan) and Reiko Akahane-Yamada (ATR Human Information Sci. Labs., Kyoto 619-0288, Japan)

Past studies on second-language (L2) speech perception have suggested that L2 learners have difficulty exploiting contextual information when perceiving L2 utterances, and that they exhibit greater difficulty than native listeners when faced with variability in temporal context. The present study investigated the extent to which native Japanese listeners, who are known to have difficulties perceiving English syllables, are influenced by changes in speaking rate when asked to count syllables in spoken English words. The stimuli consisted of a set of English words and non-words varying in syllable structure spoken at three rates by a native English speaker. The stimuli produced at the three rates were presented to native Japanese listeners in a random order. Results indicated that listeners' identification accuracy did not vary as a function of speaking rate, although it decreased significantly as the syllable structure of the stimuli became more complex. Moreover, even though speaking rate varied from trial to trial, Japanese listeners' performance did not decline compared to a condition in which the speaking rate was fixed. Theoretical and practical implications of these findings will be discussed. [Work supported by JSPS and NICT.]

2aSC28. Learning to talk: A non-imitative account of the replication of phonetics by child learners. Piers Messum (Dept. of Phonet. and Linguist., UCL, Gower St., London WC1E 6BT, UK)

How is it that an English-speaking 5-year-old comes to: pronounce the vowel of seat to be longer than that of sit, but shorter than that of seed; say a multi-word phrase with stress-timed rhythm; aspirate the /p/s of pin, polite, and spin to different degrees? These are systematic features of English, and most people believe that a child replicates them by imitation. If so, he is paying attention to phonetic detail in adult speech that is not very significant linguistically, and then making the effort to reproduce it. With all the other communicative challenges he faces, how plausible is this? An alternative, non-imitative account of the replication of these features relies on two mechanisms: (1) emulation, and (2) the conditioning of articulatory activity by the developing characteristics of speech breathing. The phenomena above then become no more than expressions of how a child finds ways to warp his phonetic output in order to reconcile conflicting production demands. The criteria he uses to do this make the challenges both of learning to talk and then of managing the interaction of complex phonetic patterns considerably more straightforward than has been imagined.

2aSC29. Acoustics of contrastive prosody in children. Rupal Patel, Jordan Piel (Dept. of Speech Lang. Pathol. & Audiol., Northeastern Univ., 360 Huntington Ave., Boston, MA 02115, r.patel@neu.edu), and Maria Grigos (New York Univ., New York, NY 10003)

Empirical data on the acoustics of prosodic control in children is limited, particularly for linguistically contrastive tasks. Twelve children aged 4, 7, and 11 years were asked to produce two utterances "Show Bob a bot" (voiced consonants) and "Show Pop a pot" (voiceless consonants) 10 times each with emphasis placed on the second word (Bob/Pop) and 10 times with emphasis placed on the last word (bot/pot). A total of 40 utterances were analyzed per child. The following acoustic measures were obtained for each word within each utterance: average fundamental frequency (f_0), peak f_0 , average intensity, peak intensity, and duration. Preliminary results suggest that 4 year olds are unable to modulate prosodic cues to signal the linguistic contrast. The 7 year olds, however, not only signaled the appropriate stress location, but did so with the most contrastive differences in f_0 , intensity, and duration, of all age groups. Prosodic differences between stressed and unstressed words were more pronounced for the utterance with voiced consonants. These findings suggest that the acoustics of linguistic prosody begin to differentiate between age 4 and 7 and may be highly influenced by changes in physiological control and flexibility that may also affect segmental features.

2aSC30. Infant-directed speech: Final syllable lengthening and rate of speech. Robyn Church, Barbara Bernhardt (School of Audiol. and Speech Sci., Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z1), Rushen Shi (Univ. of Quebec at Montreal), and Kathleen Pichora-Fuller (Univ. of Toronto)

Speech rate has been reported to be slower in infant-directed speech (IDS) than in adult-directed speech (ADS). Studies have also found phrase-final lengthening to be more exaggerated in IDS compared with ADS. In our study we asked whether the observed overall slower rate of IDS is due to exaggerated utterance-final syllable lengthening. Two mothers of preverbal English-learning infants each participated in two recording sessions, one with her child, and another with an adult friend. The results showed an overall slower rate in IDS compared to ADS. However, when utterance-final syllables were excluded from the calculation, the speech rate in IDS and ADS did not differ significantly. The duration of utterance-final syllables differed significantly for IDS versus ADS. Thus, the overall slower rate of IDS was due to the extra-long final syllable occurring in relatively short utterances. The comparable pre-final speech rate for IDS and ADS further accentuates the final syllable lengthening in IDS. As utterances in IDS are typically phrases or clauses, the particularly

strong final-lengthening cue could potentially facilitate infants' segmentation of these syntactic units. These findings are consistent with the existing evidence that pre-boundary lengthening is important in the processing of major syntactic units in English-learning infants.

2aSC31. Neural-network simulation of tonal categorization based on F_0 velocity profiles. Bruno Gauthier, Rushen Shi (Dept. of Psych., Univ. of Quebec in Montreal, C.P. 8888, Succ. Centre-ville, Montreal, QC, Canada H3C 3P8), Yi Xu (Univ. College London, London, NW1 2HE, UK), and Robert Proulx (Univ. of Quebec in Montreal, Montreal, QC, Canada H3C 3P8)

Perception studies have shown that by the age of six months, infants show particular response patterns to tones in their native language. The present study focuses on how infants might develop lexical tones in Man-

darin. F_0 is generally considered the main cue in tone perception. However, F_0 patterns in connected speech display extensive contextual variability. Since speech input to infants consists mainly of multi-word utterances, tone learning must involve processes that can effectively resolve variability. In this study we explore the Target Approximation model (Xu and Wang, 2001) which characterizes surface F_0 as asymptotic movements toward underlying pitch targets defined as simple linear functions. The model predicts that it is possible to infer underlying pitch targets from the manners of F_0 movements. Using production data of three of the speakers from Xu (1997), we trained a self-organizing neural network with both F_0 profiles and F_0 velocity profiles as input. In the testing phase, velocity profiles yielded far superior categorization than F_0 profiles. The results confirm that velocity profiles can effectively abstract away from surface variability and directly reflect underlying articulatory goals. The finding thus points to one way through which infants can successfully derive at phonetic categories from adult speech.

TUESDAY MORNING, 17 MAY 2005

GEORGIA A, 8:00 TO 11:45 A.M.

Session 2aSP

Signal Processing in Acoustics: Stochastic Signal Processing and Inversion

Max Deffenbaugh, Cochair

Exxon Mobil Upstream Research, P.O. Box 2189, Houston, TX 77252

Alan W. Meyer, Cochair

Lawrence Livermore National Lab., Univ. of California, 700 East Ave., Livermore, CA 94550

Chair's Introduction—8:00

Invited Papers

8:05

2aSP1. Signal processing techniques for inverse problems in stochastic propagation and scattering channels. Leon H. Sibul, Michael J. Roan, and Christian M. Coviello (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804-0030, lhs2@psu.edu)

The basic signal processing techniques associated with inverse problems are signal extraction, deconvolution, signal and signal parameter estimation, channel modeling and characterization. Maximum entropy, minimum cross-entropy, Kullback-Liebler divergence and other information theoretic criteria have been widely used for regularization of underdetermined inversion. We show how maximum entropy and continuous wavelet transforms can be used for spreading function (reflection density function) estimation. If the signal source or receivers are in motion through heterogeneous medium with randomly rough boundaries, propagation and scattering channels are stochastic. Stochastic channels can be characterized by stochastic Greens functions and spreading functions. If the medium probing signals are narrow band, spreading functions are random functions that show how the received signal is spread in delay and Doppler. For wideband probing signals, spreading function spread the received echo in time and time-scale dilation. If the probing signal satisfies admissibility conditions for CWT, wideband spreading functions can be estimated by inverting CWT. Rebollo-Neira and Fernandez-Rubio have shown that the continuous wavelet transform is an optimal solution of the inverse problem, estimation of the spreading function, in a maximum entropy sense. [Work supported by Office of Naval Research Code 333 and Code 321 US.]

8:30

2aSP2. Quantifying heterogeneity: Attributes, modeling, and inversion. Matthias Georg Imhof (Dept. of Geosciences, Virginia Tech, Blacksburg, VA 24061)

Characterization of reservoir heterogeneity is a necessary step in reservoir delineation, characterization, and modeling. Reservoir heterogeneity can be described deterministically, statistically, or with reservoir-forming geologic processes. Deterministic heterogeneity models are easy to build, but may have insufficient resolution and may not provide enough insights into the reservoir and its properties. The parameterization of stochastic heterogeneity models is nontrivial. Seismic data can be used to determine geostatistical parameters or to refine the geometry parameters used for Boolean modeling. The geostatistical parameters (ranges and orientations) are obtained from seismic heterogeneity attributes measuring the second-order statistics contained in small seismic datacubes. Seismic heterogeneity relates to acquisition and processing artifacts, structure, or stratigraphy and lithology. Seismic data could also be used as additional constraints in Boolean reservoir models which allows both construction of conforming reservoir models and optimization

of the geometry parameters to reduce the misfit between model and observations. Reservoirs could also be built by reconstructing their formation and evolution based on mathematical descriptions of processes such as sediment erosion, transport and deposition, compaction, deformation, subsidence and uplift, etc. The process parameters, e.g., rates of sediment input and transport, compaction, or subsidence, could be estimated by inversion of seismic and geologic data.

8:55

2aSP3. Sparseness- and continuity-constrained seismic imaging. Felix J. Herrmann (Dept. of Earth and Ocean Sci., Univ. of British Columbia, 6339 Stores Rd., Vancouver, BC, Canada V6T 1Z4, fherrmann@eos.ubc.ca)

Non-linear solution strategies to the least-squares seismic inverse-scattering problem with sparseness and continuity constraints are proposed. Our approach is designed to (i) deal with substantial amounts of additive noise ($\text{SNR} < 0$ dB); (ii) use the sparseness and locality (both in position and angle) of directional basis functions (such as curvelets and contourlets) on the model: the reflectivity; and (iii) exploit the near invariance of these basis functions under the normal operator, i.e., the scattering-followed-by-imaging operator. Signal-to-noise ratio and the continuity along the imaged reflectors are significantly enhanced by formulating the solution of the seismic inverse problem in terms of an optimization problem. During the optimization, sparseness on the basis and continuity along the reflectors are imposed by jointly minimizing the l^1 - and anisotropic diffusion/total-variation norms on the coefficients and reflectivity, respectively. [Joint work with Peyman P. Moghaddam was carried out as part of the SINBAD project, with financial support secured through ITF (the Industry Technology Facilitator) from the following organizations: BG Group, BP, ExxonMobil, and SHELL. Additional funding came from the NSERC Discovery Grants 22R81254.]

9:20

2aSP4. Reconciling data using Markov chain Monte Carlo: An application to the Yellow Sea–Korean Peninsula region. Michael Pasyanos (Lawrence Livermore Natl. Lab., 7000 East Ave., Livermore, CA 94550)

We present a technique for developing seismic models using multiple data sets using a probabilistic inverse technique. A Markov Chain Monte Carlo (MCMC) algorithm is used to sample models from a prior distribution and test them against multiple data types to generate a posterior distribution. This approach has several advantages over deterministic models, notably the reconciliation of different data types that constrain the model, the proper handling of uncertainties, and the ability to include prior information. We also benefit from the advantage of forward modeling rather than inverting the data. We demonstrate this method by using it to determine the 3-D crust and upper mantle structure for the Yellow Sea and Korean Peninsula (YSKP) region. The model is parameterized as a series of layers in a regular lat-lon grid. We start with an *a priori* model and use surface wave dispersion measurements, body wave travel times, teleseismic receiver functions, and gravity data to drive the model. The end result is a robust model distribution for the region that is best able to fit multiple data sets, and is consistent with our understanding of the tectonics of the region.

9:45

2aSP5. Global seismic tomography: Present status and future perspectives. Barbara Romanowicz (Berkeley Seismological Lab., 215 McCone Hall, Berkeley, CA 94720)

Global seismic tomography of the Earth's mantle has traditionally focused on the inversion of travel times from a small number of seismic phases, well isolated in time domain seismograms. The resulting uneven sampling of the Earth, dictated by the available distribution of earthquake sources and receivers, greatly limits the resolution of 3-D structure that can be attained. With recent progress in the simulation of wave propagation in realistic 3-D Earth models combined with significant increases in computational power, seismologists are well positioned to start fully exploiting the wealth of information contained in the rapidly expanding database of high quality broadband seismic waveforms, accumulated globally and regionally over the last twenty five years. New opportunities presented by 3-D forward and inverse modeling of long period waveforms will be presented with illustrations from recent work on deep and shallow elastic and anelastic mantle structure, addressing, among others, such questions as: what is the fate of subducted slabs or the nature of the lower mantle "superplumes"?

10:10–10:30 Break

10:30

2aSP6. Surface wave tomography from seismic ambient noise in Southern California. Karim Sabra, Peter Gerstoft, Philippe Roux, William Kuperman (Marine Physical Lab., Scripps Inst. of Ocean., UCSD, La Jolla, CA 92093-0238, ksabra@mpl.ucsd.edu), and Michael Fehler (Los Alamos Natl. Lab., NM)

It has been demonstrated experimentally that an estimate of the Green's tensor between two seismic stations can be obtained from the long-time average of the cross-correlation of ambient noise at the two stations. This result provides a means to image Earth structure using the ambient noise field only, without the use of active seismic sources or earthquakes. Seismic noise data from 148 broadband seismic stations in Southern California were used to extract the surface wave arrival-times between all station pairs in the network in the frequency band 0.05–0.4 Hz. In this frequency band, ambient noise (originating from ocean microseisms) propagating over long distances is typically dominated by surface waves. A record section of the waveforms as a function of increasing receiver separation shows clearly that the recovered signals are propagating wavetrains. The seismic data were then used in a simple, but densely sampled tomographic procedure to estimate the surface wave velocity structure for a region in Southern California. The result compares favorably with previous estimates obtained using more conventional and elaborate inversion procedures. This demonstrates that coherent ambient noise between station pairs can be used for seismic imaging purposes.

10:55

2aSP7. An overview of geophysical investigations at archaeological sites in Helike, Greece. Alan Witten and Jamie Rich (School of Geology and Geophys., Univ. of Oklahoma, 810 Energy Ctr., Norman, OK 73019)

Helike is an area on the south-central coast of the Gulf of Corinth. This name is derived from the lost Classical Greek city of Helike that vanished during an earthquake and subsequent tsunami in 373 BC. It was originally believed that the remains of Helike were located offshore in the Gulf of Corinth; however, early searches of this area proved fruitless. More recently, boreholes developed in the coastal plain revealed marine sediments that date to the Classical Greek period suggesting that, as a result of uplift and/or sedimentation, Helike now lies onshore. This find stimulated a sequence of ongoing geophysical investigations and excavations that have revealed buried architecture and artifacts from three periods; Early Bronze Age, Classical Greek, and Roman.

11:20

2aSP8. Acoustic archeology at the lost city of Helike. Sean K. Lehman (Lawrence Livermore Natl. Lab., L-154, 7000 East Ave., Livermore, CA 94550)

In 373 BC an earthquake and tsunami destroyed and submerged the classical Greek city of Helike (HELL-E-KEY) on the north Peloponnese shore of the Gulf of Corinth. In June 2004, surface seismo-acoustic surveys were performed in the area of present Helike in order to collect non-invasive tomographic data to guide the archeological excavations. Prior to tomographic inversion and imaging, the raw data time series must be preprocessed to estimate ground acoustic velocity, and to remove the ground roll signal. We present a generalized cross-correlation technique which appears successful in solving both problems. We describe the technique as applied to the data and present tomographic inversions.

TUESDAY MORNING, 17 MAY 2005

GEORGIA B, 7:40 A.M. TO 12:00 NOON

Session 2aUW

Underwater Acoustics and Acoustical Oceanography: Acoustic Interaction with Ocean Boundaries: Single Bounce Measurements

Charles W. Holland, Chair

Applied Research Lab., Pennsylvania State Univ., State College, PA 16804-0030

Chair's Introduction—7:40

Contributed Papers

7:45

2aUW1. Direct-path rough surface scattering experiments with milled surfaces in acoustic tank facilities. Raymond J. Soukup, Robert F. Gragg, Jason E. Summers, and Edward L. Kunz (Naval Res. Lab., 4555 Overlook Ave., Washington, DC 20375, soukup@abyss.nrl.navy.mil)

To investigate specific issues relating to rough surface scattering theoretical predictions, the Naval Research Laboratory has been conducting scaled-down versions of its direct-path ocean bottom scattering experiments. These experiments utilize milled elastic rough surfaces in acoustic tank facilities. Investigations into the dependence of scattering on roughness and geoacoustic parameters lend themselves to a systematic treatment with milled surfaces that are fabricated to match specific roughness spectral parameters. The characterization and fabrication of such rough surfaces, and the methodology for creating experimental geometries that are analogous to previous measurements with elastic ocean bottoms, are described. A set of scattering strength measurements in the 100–300 kHz band with a quasi-monostatic geometry is compared with the theoretical predictions given by perturbation theory and the NRL small slope model. The results show that these experiments are a viable means of verifying the predictions of the scattering models, given a sufficiently large statistical sample. [Work supported by ONR.]

8:00

2aUW2. Observations of a geocloud feature in the straits of Sicily. Thomas C. Weber, Charles W. Holland (Appl. Res. Lab., The Pennsylvania State Univ., State College, PA 16804), and Giuseppe Etiope (Natl. Inst. of Geophys. and Volcanology, Rome, Italy)

Several persistent sonar clutter features were observed in 2002 during an experiment in the Straits of Sicily (Malta Plateau) in a region that is nominally flat and thickly sedimented. High frequency sidescan and seismic reflection data from the region indicated that mud volcanoes were present and were possibly expelling clouds of methane bubbles. In 2004 the region was revisited with oceanographic equipment including a multi-beam sonar, an acoustic Doppler current profiler (ADCP), CTD's, temperature loggers, a methane sensor, and a video camera. The visual observations at one site show carbonate mounds comprised of large heterogeneous blocks that suggest the presence of mud volcanism. A multibeam sonar survey of the area shows that the carbonate mounds are of order 5–10 m in height and 10–100 m in lateral dimension, and that they form in clusters. The multibeam backscatter data show high amplitudes in the region surrounding the carbonate mounds, possibly indicating the surficial extent of the carbonate material. High levels of methane were observed in the water column above the mounds, although conclusive evidence of the presence of methane bubble clouds was not found. However, high frequency backscatter recorded on the ADCP above multiple

carbonate mounds showed plume-like features that may be bubble clouds, suspended fine-grained sediment, and/or schools of fish.

8:15

2aUW3. Close-range acoustic scattering from mud volcanoes.

Charles W. Holland, Thomas C. Weber (Appl. Res. Lab., Penn State Univ., State College, PA, 16804), and Giuseppe Etiope (Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy)

Submarine mud volcanoes occur in many parts of the world's oceans and form an aperture for gas (mostly methane) and fluidized mud emission from the earth's interior. Their characteristics are of considerable interest to the geology, geophysics, geochemistry, and underwater acoustics communities. For the later community, mud volcanoes are important because they pose a potential source of clutter for active sonar. Close-range (single boundary interaction) scattering measurements from a mud volcano in the Straits of Sicily were conducted with a vertical source and receive array. The data show target strengths from 800–3600 Hz of 6–12 dB for a monostatic geometry with grazing angles of 3–5 degrees. The target strengths are very similar for vertically bi-static paths with incident grazing angles of 3–5 degrees and scattered angles of 33–45 degrees. The evidence suggests that the scattering mechanism is the mud volcano (carbonate) structure. [Work supported by the Office of Naval Research and NATO Undersea Research Centre.]

8:30

2aUW4. Comparison of the magnitude and phase of the reflection coefficient from a smooth water/sand interface with elastic and poroelastic models.

Marcia Isakson (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029), H. John Camin (The Penn State Univ., State College, PA 16804-0030), and Gaetano Canepa (NATO Undersea Res. Ctr., La Spezia, Italy)

The reflection coefficient from a sand/water interface is an important parameter in modeling the acoustics of littoral environments. Many models have been advanced to describe the influence of the sediment parameters and interface roughness parameters on the reflection coefficient. In this study, the magnitude and phase of the reflection coefficient from 30 to 160 kHz is measured in a bistatic experiment on a smoothed water/sand interface at grazing angles from 5 to 75 degrees. The measured complex reflection coefficient is compared with the fluid model, the elastic model and poro-elastic models. Effects of rough surface scattering are investigated using the Bottom Response from Inhomogeneities and Surface using Small Slope Approximation (BoRIS-SSA). Spherical wave effects are modeled using plane wave decomposition. Models are considered for their ability to predict the measured results using realistic parameters. [Work supported by ONR, Ocean Acoustics.]

8:45

2aUW5. Laboratory study of high frequency scattering from water-saturated sandy sediments.

Anatoliy Ivakin (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, ivakin@apl.washington.edu), Jean-Pierre Sessarego, and Patric Sanchez (CNRS/LMA, Marseille cedex 20, France)

Backscattering from a water-saturated well-sorted fine sand and from this sand with various inclusions was studied in a laboratory tank at frequencies from 200 kHz to 3 MHz and grazing angles from 20 to 90 degrees. Acoustic and granulometric properties of the sand are presented and discussed in a companion paper. First, the measurements of the frequency-angular dependencies of the backscattering strength were made for a thick sediment layer with a smoothed surface and without any inclusions in the volume. Then glass beads and coarse sand particles were used as additional scatterers or inclusions placed on the sediment surface or covered by sediment layers of various thickness. The data were compared with a previously reported model of scattering from discrete inclusions in the sediment [A. N. Ivakin, *J. Acoust. Soc. Am.* **116**, 2575 (2004)]. The

model inputs are material parameters of the sediment and inclusions and the size-depth distribution of inclusions. The results of model/data comparison are discussed and possibilities for inversion of various sediment parameters from backscattering data are shown.

9:00

2aUW6. Theory and experiments for surface-wave focused acoustic arrivals: A deterministic view.

Grant Deane (Mail Code 0238, Scripps Inst. of Oceanogr., UCSD, La Jolla, CA 92093-0238) and James Preisig (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Statistical descriptions of surface wave scattering are now highly developed, and have proven very effective in relating the first and second order statistics of surface-reflected acoustic arrivals to surface conditions. Notwithstanding this success, the deterministic features of surface-scattered arrivals are of considerable importance. The limits of performance of underwater communications systems in very shallow water, for example, are determined by the frequency and properties of extreme surface focusing events. The analysis of data from the Wavefronts I, II, and V series of experiments has led to an understanding of the wave properties controlling the focusing events, as well as their impact on acoustic channel estimators. Data from the wavefronts experiments (SIO pier, 7 m deep, 40 m range) and the SPACE02 experiment (Martha's Vineyard, 15 m depth, 250–1000 m range) will be presented and discussed. [Work supported by ONR.]

9:15

2aUW7. Forward scattering from the sea surface: Observations of both subtle and profound effects of bubbles in single-interaction measurements.

Peter H. Dahl (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

For frequencies of O(10) kHz and above, field data show that near-surface bubbles impact forward scattering from the sea surface in three phases. The first occurs under mild conditions (wind speed less than 5–7 m/s); here a pulse forward scattered from the sea surface is extended in time, but only at levels ~30 dB below the peak level, which itself is not attenuated. The second occurs under more vigorous conditions (wind speed 7–12 m/s); here a significant energy loss is observed, but time and angle spreading (dominated by rough surface scattering) remain relatively unchanged. The third occurs under still more vigorous conditions (wind speed greater than ~12 m/s). Here, there is near total occlusion of the sea surface, time and angle spreading are manifestly altered, and bubble-mediated energy loss becomes bounded by scattering from bubbles. Examples from ASIAEX East China Sea and other archival data sets will be discussed along with a model for bubble-mediated energy loss in forward scattering from the sea surface. In the case of near total occlusion, an interesting example of the knock-down of horizontal coherence will be discussed along with a model that utilizes the van Cittert-Zernike Theorem. [Research supported by ONR Ocean Acoustics.]

9:30

2aUW8. On the relationship between sea state and the coherent-to-incoherent intensity ratio for high frequency, shallow water propagation.

Philippe Roux, William A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., UCSD, 8820 Shellback Way, La Jolla, CA 92093-0238), R. Lee Culver, Steven D. Lutz, David L. Bradley (State College, PA 16804, rlc5@psu.edu), and Mark Stevenson (NATO Undersea Res. Ctr., La Spezia, Italy)

A 29-element source array and 32-element receiver array have been deployed in shallow water to investigate various aspects of the acoustic channel. Both arrays spanned much of the water column. Source-receiver range was 800 m, and 3.5 kHz pulses were transmitted using one source at a time and recorded on all elements. In this analysis, we compare the ratio of coherently to incoherently averaged intensities for each source-receiver pair, and averaged over all such pairs, to measure surface roughness. We find that the coherent-to-incoherent intensity ratio decreases with time after the direct path arrival, and that the slope becomes more negative as

wind speed and thus surface roughness increases. A simple model for signal phase variation induced by surface scatter shows a direct relationship between the degree of phase variation and wind speed. The relationship between the slope of the coherent-to-incoherent intensity ratio and a physics-based model of incoherent sea surface forward scatter is also considered. [Work supported by ONR Code 321US.]

9:45

2aUW9. Scintillation index of ocean surface forward scattered HF acoustic signals: Beam pattern and pulse length effects. Benjamin Cotté, David Bradley, and R. Lee Culver (Appl. Res. Lab., Penn State Univ., State College, PA 16804, bzc125@psu.edu)

The intensity fluctuations of surface bounce path measured during a field test near San Clemente Island in August 2002 have been presented previously. Signals used during this test were 20- and 40-kHz continuous wave (CW) pulses with 0.14- and 1.0-ms durations (“short” and “long” pulses, respectively). Experimental results have been compared to predictions of a model by Yang and McDaniel [Waves Random Media, **1**, 419–439 (1991); J. Acoust. Soc. Am. **91**, 1960–1966 (1992)]. Their initial model assumes CW signals and omni-directional projectors and receivers. We observed a good fit between model and experimental results for the long pulse data when the specular point was in the main lobe of both projector and receiver. The discrepancies observed in the other cases were attributed to beam pattern and pulse length effects. We have extended the model to incorporate these two effects. [Work supported by ONR Code 321US.]

10:00–10:15 Break

10:15

2aUW10. Measurements of mid-frequency bottom backscattering strengths on the Outer Hebrides Platform and the Stanton Banks. Edward L. Kunz and Roger C. Gauss (Nav. Res. Lab, Code 7140, Washington, DC 20375-5350, edward.kunz@nrl.navy.mil)

In July 2002, direct-path, mid-frequency (2–5 kHz) bottom backscattering strength measurements were conducted at 21 sites on the Outer Hebrides Platform and the Stanton Banks (northwest of Scotland) during The Technical Cooperation Program’s T-MAST 02 sea trial. This is a complex area, with bottom grabs indicating the wide presence of stones, pebbles, gravel and rocks throughout most of the region. When stones, etc. were present, bottom backscattering strengths were generally high and flattish in angle (over 10 to 40 deg), with many sites empirically exhibiting a $\sin\theta$ or flatter angular dependence. The frequency dependence in these cases was mild. When stones, etc. were absent, bottom backscattering strengths were less high and less flat in angle, with strong frequency dependence. These results dramatically illustrate the inadequacies of using Lambert’s Law to model bottom backscattering strengths and suggest that active sonars operating in this environment will encounter strong reverberation. [Work supported by ONR.]

10:30

2aUW11. A semi-empirical model for predicting surface scattering strengths. Roger C. Gauss, Joseph M. Fialkowski, and Daniel Wurmser (Nav. Res. Lab, Code 7144, Washington, DC 20375-5350, roger.gauss@nrl.navy.mil)

Accurate monostatic and bistatic active sonar performance predictions require accurate predictions of reverberation which in turn rely on accurate estimates of surface scattering strength. This paper presents a new model that estimates the surface scattering strength due to both the rough air-sea interface (small-slope theory) and subsurface bubble clouds (a Lloyd-mirror model), given the acoustic frequency (<5 kHz), incident grazing angle, scattered grazing angle, bistatic angle, wind speed, and sound speed in the water. This semi-empirical surface scattering strength (SESS) model advances previous incarnations [Gauss and Fialkowski, J. Acoust. Soc. Am. **105**, 1254 (1999) and ECUA 2000] by: (1) incorporating a low wavenumber cut-off to the surface-wave roughness spectra; (2) adding

significant sets of low- and mid-frequency, open-ocean backscattering strength data to the database used to derive model parameters (using more comprehensive procedures); and (3) adding a new angle-dependent factor to the bubble-cloud scattering formula. All these changes have increased prediction accuracy, with variances of least-square fits to the new, full data set down from 9.7 dB² for the old model to 5.3 dB² for the new model. [Work supported by ONR and SPAWAR PMW-180.]

10:45

2aUW12. Backscattering of short acoustic pulses from 3-D rough surfaces: Statistical properties of first arrivals. Oleg A. Godin (CIRES, Univ. of Colorado and NOAA/Environ. Technol. Lab., 325 Broadway, Boulder, CO 80305), Isosif M. Fuks, and Mikhail I. Charnotskii (Zel Technologies, Broadway, Boulder, CO 80305)

Time history of a pulse backscattered by a rough surface contains information about position of the surface and properties of the scatterers. Such information is utilized successfully in a number of remote sensing techniques ranging from echo sounding of the ocean bottom to medical ultrasonics and satellite altimetry. In this paper, statistical properties of backscattered waves are considered in the geometrical acoustics approximation. The probing pulse duration is assumed to be sufficiently short so that signals backscattered in a vicinity of individual specular points on a rough surface do not overlap in time. Theoretical results previously obtained in a 2-D problem [I. M. Fuks and O. A. Godin, Waves Random Media **14**, 539–562 (2004)] are extended to sound scattering from 3-D rough surfaces by following an approach developed in stochastic geometry. Predictions of an asymptotic theory are verified against results of a numerical simulation. Travel times and intensities of the first and the second arrivals of the backscattered wave are quantified in terms of statistical moments of roughness. It is found that, as in 2-D case, the travel time and the intensity are strongly correlated; on average, the earlier a signal arrives, the smaller is its intensity.

11:00

2aUW13. Backscattering from targets residing in caustics resulting from ocean boundary interactions. Benjamin R. Dzikowicz (Naval Surface Warfare Ctr., Panama City, FL 32407) and Philip L. Marston (Washington State Univ., Pullman, WA 99164-2814)

Detection of targets by backscatter in shallow water can be enhanced by interactions with ocean boundaries. A laboratory experiment is performed where a spherical target passes through an Airy caustic formed by a curved surface. When the target resides in the insonified region of the caustic there are two sets of multi-path rays: two pairs reflecting once off the surface (either to or from the target), and three reflecting twice off the surface (to and from the target). When a target moves across the caustic the singly reflected rays merge, as do the doubly reflected. With a longer tone burst the rays in each set overlap and the backscatter is greatly enhanced as the target moves into the insonified region. For a point target the singly reflected backscatter scales as an Airy function [B. R. Dzikowicz and P. L. Marston, J. Acoust. Soc. Am. **116**, 2751–2757 (2004)], and the doubly reflected as the square of an Airy function. For a finite target the doubly reflected backscatter unfolds into a hyperbolic umbilic function. The arguments of the Airy and Hyperbolic Umbilic functions are calculated using the relative echo times of transient pulses. [Work supported by ONR.]

11:15

2aUW14. Determining flat interface reflection coefficients using forward scattering from a rippled sediment interface. Kevin L. Williams (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

For sediment interfaces that are very rough on the scale of the acoustic wavelength (i.e., kh greater than 1 where k is $2\pi/\text{wavelength}$ and h is the rms roughness of the water/sediment interface) it is possible to estimate what the reflection coefficient would be if the interface were flat. In order to do so, a large ensemble of forward scattering measurements are

needed in order to reduce the statistical uncertainty of the estimated reflection coefficient. In addition to the statistical uncertainty there can be biases in the estimate (for some grazing angles) that must be taken into account. The above conclusions will be supported through the use of monte carlo simulations of scattering from a rippled interface. The simulations are carried out in the context of discriminating between alternative acoustic models of sand sediments. [Work supported by ONR.]

11:30

2aUW15. Bayesian inversion of frequency-averaged reflection data. Stan E. Dosso (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC, Canada) and Charles W. Holland (Penn State Univ., State College, PA 16801)

This paper presents a nonlinear Bayesian inversion of high-resolution seabed reflection data to estimate visco-elastic parameters of the upper sediments. The inversion is applied to data from two sites in the Strait of Sicily. One site is characterized by low-velocity, silty-clay sediments, resulting in data with a well-defined angle of intromission. The second site is characterized by high-velocity clayey sand, resulting in a critical angle. The data were frequency averaged from 500–2000 Hz and inverted for visco-elastic parameters of a half-space seabed model. The likelihood function employed in the inversion is based on the assumption of independent, Gaussian-distributed data errors, with the standard deviation included as a nuisance parameter in the inversion. Statistical tests are applied to the data residuals *a posteriori* to validate these assumptions. Good results are obtained for sediment compressional-wave velocity, compressional attenuation, and density; shear parameters are less-well determined although low shear-wave velocities are indicated. The Bayesian analysis

provides a quantitative comparison of inversion results for the two sites in terms of the resolution of specific geoacoustic parameters, and indicates that the geoacoustic information content is significantly higher for angle-of-intromission data.

11:45

2aUW16. Bayesian inversion of multi-frequency reflection data with strongly correlated errors for density gradients. Jan Dettmer, Stan E. Dosso (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC, Canada), and Charles W. Holland (Penn State Univ., State College, PA 16801)

This paper develops a non-linear Bayesian inversion for multi-frequency reflection-loss data with strongly correlated data errors to resolve density and sound-velocity gradients which are often observed in the uppermost sediment layer. Although data errors are usually assumed to be independent in geoacoustic inversion, in reality measured data often show strong error correlations. The inversion developed here is designed to take error correlations into account. A full data covariance matrix is estimated from initial residuals of non-uniformly sampled data. This covariance matrix is then used in the likelihood function of a fast Gibbs sampler to sample the posterior probability density and provide parameter estimates and credibility intervals. Rigorous statistical tests are applied to the resulting data residuals to illustrate the benefits of this error treatment. The approach is applied to reflectivity data collected at a site characterized by low-velocity, water-saturated sediments in the Strait of Sicily. Density and sound-speed gradients are clearly resolved by the reflectivity data and agree with core measurements from the experiment site within the credibility bounds.

Meeting of the Standards Committee Plenary Group

to be held jointly with the

ANSI-Accredited U.S. Technical Advisory Group (TAG) Meetings for:

ISO/TC 43 Acoustics
ISO/TC 43/SC 1 Noise and
IEC/TC 29 Electroacoustics

P. D. Schomer, Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 43 Acoustics and ISO/TC 43/SC 1 Noise
2117 Robert Drive, Champaign, IL 61821

V. Nedzelnitsky, U.S. Technical Advisor (TA) for IEC/TC 29, Electroacoustics
*National Institute of Standards and Technology (NIST), Sound Building, Room A147, 100 Bureau Drive, Stop 8221,
Gaithersburg, MD 20899-8221*

The meeting of the Standards Committee Plenary Group will precede the meetings of the Accredited Standards Committees S1, S2, S3, and S12, which are scheduled to take place in the following sequence:

ASC S1 Acoustics	17 May 2005	1:45 p.m. to 3:15 p.m.
ASC S12 Noise	17 May 2005	3:30 p.m. to 5:00 p.m.
ASC S2 Mechanical Vibration and Shock and TAGs to TC 108	18 May 2005	8:00 a.m. to 9:30 a.m.
ASC S3 Bioacoustics	18 May 2005	10:30 a.m. to 12:00 noon

Discussion at the Standards Committee Plenary Group meeting will consist of national items relevant to all S Committees.

The ANSI-Accredited US Technical Advisory Group (TAGs) for ISO/TC 43 Acoustics and IEC/TC 29 Electroacoustics, whose membership consists of members of S1 and S3, and other persons not necessarily members of these Committees, will meet during the Standards Plenary meeting. The ANSI-Accredited US Technical Advisory Group (TAG) for ISO/TC 43/SC 1 Noise, whose membership consists of the members of S12 and other persons not necessarily members of S12, will meet as well. The reports of the Chairs of these TAGs will not be presented at any other S Committee meeting. There will be a report on the interface of S1 and S3 activities with those of ISO/TC 43 and IEC/TC 29 including plans for future meetings of ISO/TC 43 and IEC/TC 29.

Members of S2 Mechanical Vibration and Shock (and U.S. TAG for ISO/TC 108 and its Subcommittees, (SC2, SC3, SC4, SC5, and SC6) are also encouraged to attend the Standards Committee Plenary Group meeting even though the S2 meeting will take place the next day.

The U.S. Technical Advisory Group (TAG) Chairs for the various international Technical Committees and Subcommittees under ISO and IEC, which are parallel to S1, S2, S3, and S12 are as follows:

<u>U.S. TAG Chair/Vice Chair</u>	<u>TC or SC</u>	<u>U.S. TAG</u>
ISO		
P. D. Schomer, Chair	ISO/TC 43 Acoustics	S1 and S3
P. D. Schomer, Chair	ISO/TC 43/SC1 Noise	S12
D. J. Evans, Chair	ISO/TC 108 Mechanical Vibration and Shock	S2
A. F. Kilcullen, Co-Chair R. Taddeo, Co-Chair	ISO/TC 108/SC2 Measurement and Evaluation of Mechanical Vibration and Shock as Applied to Machines, Vehicles and Structures	S2
D. J. Evans, Chair	ISO/TC 108/SC3 Use and Calibration of Vibration and Shock Measuring Instruments	S2
D. D. Reynolds, Chair	ISO/TC 108/SC4 Human Exposure to Mechanical Vibration and Shock	S3
D. J. Vendittis, Chair R. F. Taddeo, Vice Chair	ISO/TC 108/SC5 Condition Monitoring and Diagnostics of Machines	S2
G. Booth, Chair	ISO/TC 108/SC6 Vibration and Shock Generating Systems	S2
IEC		
V. Nedzelnitsky, U.S. TA	IEC/TC 29 Electroacoustics	S1 and S3

Session 2pAAa

**Architectural Acoustics, Education in Acoustics, Noise and Psychological and Physiological Acoustics:
Topical Meeting on Classroom Acoustics—The Research Perspective III**

Murray R. Hodgson, Cochair

*Univ. of British Columbia, School of Occupational and Environmental Hygiene, 2206 East Mall,
Vancouver, BC V6T 1Z3, Canada*

Lily M. Wang, Cochair

*Univ. of Nebraska Lincoln, Architectural Engineering, Peter Kiewit Institute, 1110 South 67th St., Omaha, NE 68182-0681***Chair's Introduction—1:10****Invited Papers****1:15**

2pAAa1. Subjective evaluation of speech and noise in learning environments in the realm of classroom acoustics: Results from laboratory and field experiments. Markus Meis (Hearing Ctr. Oldenburg/Univ. of Oldenburg, Marie-Curie-Str. 2, 26129 Oldenburg, Germany, markus.meis@hoerzentrum-oldenburg.de), Christian Nocke (Akustikbuero Oldenburg), Simone Hofmann (Freudenberg Bldg. Systems), and Bernhard Becker (Deutsche Rockwool Mineralwool GmbH & CO. OHG)

The impact of different acoustical conditions in learning environments on noise annoyance and the evaluation of speech quality were tested in a series of three experiments. In Experiment 1 ($n=79$) the auralization of seven classrooms with reverberation times from 0.55 to 3.21 s [average between 250 Hz to 2 kHz] served to develop a Semantic Differential, evaluating a simulated teacher's voice. Four factors were found: acoustical comfort, roughness, sharpness, and loudness. In Experiment 2, the effects of two classroom renovations were examined from a holistic perspective. The rooms were treated acoustically with acoustic ceilings (RT=0.5 s [250 Hz–2 kHz]) and muffling floor materials as well as non-acoustically with a new lighting system and color design. The results indicate that pupils ($n=61$) in renovated classrooms judged the simulated voice more positively, were less annoyed from the noise in classrooms, and were more motivated to participate in the lessons. In Experiment 3 the sound environments from six different lecture rooms (RT=0.8 to 1.39 s [250 Hz–2 kHz]) in two Universities of Oldenburg were evaluated by 321 students during the lectures. Evidence found supports the assumption that acoustical comfort in rooms is dependent on frequency for rooms with higher reverberation times.

1:45

2pAAa2. Prevalence, nature and risks of voice problems among public school teachers. Linda Rammage (Dept. of Surgery, UBC, PVCPRP, 4th Fl. Willow Pavilion, VGH, 805 West 12th Ave, Vancouver, BC, Canada V5Z 1M9, lira@interchange.ubc.ca), Murray Hodgson (Univ. of British Columbia, Vancouver, BC V5Z 1M9), and Charlie Naylor (British Columbia Teachers Federation)

Voice problems among teachers represent a rising cause of teacher absenteeism, use of sick benefits, and stress among teachers and students. In British Columbia, the BC Teachers Federation and Workers Compensation Board are receiving increasing numbers of claims from teachers experiencing occupational voice problems and in the provincial voice clinic, the percentage of teachers in the clinic population is rising. Previous studies of teachers voice problems have typically had low return rates, which can bias the prevalence estimates, and have not incorporated standardized voice inventories, psychological inventories and acoustic measures. A survey study is in progress in B.C. to probe demographic, environmental, voice-use, health, psychological and personality issues that are thought to contribute to development of voice problems among teachers. To ensure validity of prevalence estimates by high return rates, on-site completion of questionnaires is being used in schools. Acoustical measures are also being made of representative classrooms, to determine the degree to which noise and reverberation contribute to voice problems among teachers.

2:05

2pAAa3. Benefits of teaching voice amplification as related to subjective laryngeal symptoms and perceived voice quality in teachers. Valdis Jonsdottir (Hraukbajarkot v/ Akureyri 601, Iceland)

Loud speaking due to noisy working conditions is a common cause for teachers' voice disorders. One way to diminish the vocal load of teaching is to make use of technical equipment. This Icelandic study explores: (1) if the use of amplification in classrooms would diminish the teachers' experienced symptoms of vocal fatigue; and (2) whether there is a possible change in perceptual voice quality during a teachers' working day. Thirty-three teachers, from grade school to university level, voluntarily served as subjects. They used amplifiers while teaching for one week at least. After that, they filled out a questionnaire concerning their symptoms and experiences. The results showed that the majority of teachers found amplification beneficial. They found it easier to talk and experienced less fatigue. The few disadvantages were technical. For a perceptual analysis, three females and two males (mean age 51 years) with long teaching experience and three or more dysphonic symptoms during the term, had their speech recorded while teaching, with and without amplification. In the clinical examination, no pathological changes were found in the vocal folds. In both studies, the quality of the voices was esteemed better when amplification was used.

2pAAa4. Benefits of teaching voice amplification as related to subjective laryngeal symptoms in teachers, and to the listening conditions for pupils. Valdis Jonsdottir (Hraukbjarkot v/ Akureyri 601, Iceland)

Teachers unfavorable working conditions—large classrooms with poor acoustics, background noise and the great distance between teacher and pupils often are the main reasons for teachers voice problems. For pupils, the same circumstances simultaneously create unfavorable listening conditions. One way to diminish the vocal load of teaching is by making use of technical equipment. This Icelandic study explores whether voice amplification usage: (1) diminishes the teachers subjective symptoms of vocal fatigue; (2) makes it easier for pupils to listen and follow lessons; and (3) reveals other benefits and/or disadvantages of the usage of amplification in classrooms. Thirty-three teachers with their students (total 791), from grade school to university level, volunteered as subjects. The teachers used amplification in class for at least a week. At the end of the research period, the participants answered questionnaires, except the pupils between 6–9 years of age who were asked two questions. The results showed: (1) the majority of teachers found it easier to talk and experienced less fatigue using vocal amplification; (2) the students found it easier to hear the teacher through class chatter and to follow the lessons; (3) the disadvantages mentioned were mainly technical problems.

2pAAa5. Acoustic and other factors relating to the use of sound field systems in classrooms. Anne Carey, Bridget Shield (Dept. of Eng. Systems, Faculty of Eng., Sci. and Built Environment, London South Bank Univ., London SE1 0AA, UK), Julie Dockrell, and Kate Rigby (Univ. of London, 25 Woburn Square, London WC1H 0AA, UK)

A study has been made of thirty-five primary school classrooms which have Sound Field Systems (SFS) installed. Acoustic surveys of the classrooms have been undertaken and detailed observations made of the physical characteristics of each room such as its construction and design, plus any acoustic treatment. The positioning and type, e.g., column or separate speakers, of SFS installed were also noted. Details of the procedures for purchasing, installing, and maintaining the SFS were obtained where possible, together with information concerning training of teachers in their use. Results from these surveys suggest that in many cases SFS are installed in rooms where their performance may be compromised because of inadequate acoustic conditions. Additionally problems may arise due to inappropriate installation, choice of system, poor maintenance, or lack of training. The effectiveness of acoustic treatment such as absorption applied to ceilings and walls has also been investigated in a range of classrooms and its impact on the use of SFS considered.

3:05–3:20 Break

2pAAa6. Installation and impact of sound field systems on hearing and hearing impaired children and their teachers. Julie Dockrell, Kate Rigby (School of Psych. and Human Development, Inst. of Education, Univ. of London, 25 Woburn Square, London WC1H 0AA, UK), Bridget Shield, and Anne Carey (London South Bank Univ., London SE1 0AA, UK)

An evaluation of the installation and use of sound field systems in ten schools in England has been carried out. The evaluation included noise surveys of classrooms, questionnaire surveys of pupils and teachers and experimental testing of children with and without the use of SFS. The aim of this project was to investigate the impact of SFS on teaching and learning in elementary school classrooms, in particular, to ascertain whether the SFS differentially benefited children with hearing impairments. Barriers to teachers use of SFS were found in terms of equipment placement and maintenance, appropriate training, and teacher's knowledge. Nonetheless positive reports are recorded from both teachers and pupils. Teachers' and pupils' perceptions are compared with objective data evaluating change in performance when SFS are used for language and cognitive tasks. Data from children with hearing impairments and additional learning needs are analyzed for comparative purposes. The results are discussed in terms of effective practice for the use of SFS with elementary school pupils.

2pAAa7. Acoustic and electromagnetic noise from lighting in classrooms. Charles A. Laszlo, Jonathan Lashin (UBC Dept. of Elec. and Comput. Eng., 2356 Main Mall, Vancouver, BC, Canada V6T 1Z4, claszlo@telus.net), and Murray R. Hodgson (UBC Acoust. and Noise Res. Group, Vancouver, BC, Canada V6T 1Z3)

Following complaints by hard-of-hearing students using assistive-listening devices, and their teachers, the hum-like noise generated by fluorescent lighting was investigated in classrooms and the school library in a typical school. This hum is caused by vibrations in the core of the magnetic ballasts. Measurements were made in several rooms without students present. Noise levels increased between 7 and 15 dB when fixtures using magnetic ballasts were switched on. Spectral analysis showed the presence of 30, 60, 120, and 240 Hz components. In rooms where electronic ballasts were installed, there was no increase in noise level when the lights were switched on. Since hearing aids and assistive-listening devices worn by students may also be influenced by magnetic fields, these were also surveyed in these classrooms. The magnetic fields generated by the lights were not significant, but near some wiring and electrical panels the interference was strong. In rooms with electronic ballasts some infrared assistive-listening devices picked up strong high-frequency hum. It is recommended that the effect of lighting fixtures and the electrical-distribution system be taken into account in the acoustical and communication design of classrooms.

4:00

2pAAa8. Real world issues in classroom acoustics for hearing impaired students. Klaus Kleinschmidt (Acoust. Consultant, 132 Mary Catherine Dr., Lancaster, MA 01523)

Experience is being gained from evaluating and modifying the acoustical properties and background noise levels of existing classrooms for use by hearing impaired students. Projects include more than 25 schools in public school systems that are mainstreaming handicapped children. Various practical and economic restrictions have led to recommendations for modifications that do not necessarily comply with ANSI S12.60 2002. Examples of real world conditions and practical solutions will be described.

4:15

2pAAa9. Noise problem in a primary level classroom. Sergio Beristain (IMA, ESIME, IPN. P.O. Box 12-1022, Narvarte, 03020, Mexico D. F., Mexico. sberista@hotmail.com)

Noise was assessed in a primary school in Mexico City (six to twelve year old kids), known as having significant noise inside classrooms. Several sources were identified, among them, traffic noise and noise from the sports and recreational areas, which were in the center of the school with the classrooms surrounding them. Reverberation was an issue, but not very significant. But in some cases the most annoying sound (noise), was the one made by the students within the classroom, inducing a number of reactions in the lecturers. These reactions ranged from shouting to give the lecture, through trying to control the students, to ignore them and let it go. Lecturer's voice levels were also measured, finding out a wide spread in normal speech voice level, and some of their experiences and comments were analyzed. Results are summarized in this presentation.

TUESDAY AFTERNOON, 17 MAY 2005

BALMORAL, 1:30 TO 5:00 P.M.

Session 2pAAb

Architectural Acoustics and the National Council of Acoustical Consultants: Student Design Competition

Robert C. Coffeen, Cochair

Univ. of Kansas, School of Architectural and Urban Design, Marvin Hall, Lawrence, KS 66045

Byron Harrison, Cochair

The Talaske Group Inc., 105 N Oak Park Ave., Oak Park, IL 60301

The Technical Committee on Architectural Acoustics of the Acoustical Society of America and the National Council of Acoustical Consultants are sponsoring this Student Design Competition that will be professionally judged at this meeting. The purpose of this design competition is to encourage students enrolled in architecture, architectural engineering, and other university curriculums that involve building design and/or acoustics to express their knowledge of architectural acoustics in the design of a drama theater complex located within an urban mixed-use development. This competition is open to undergraduate and graduate students from all nations.

Submissions will be poster presentations that demonstrate room acoustics, noise control, and acoustic isolation techniques in building planning and room design.

The submitted designs will be displayed in this session and they will be judged by a panel of professional architects and acoustical consultants. An award of \$1,250 US will be made to the submitters of the entry judged "First Honors." Up to four awards of \$700 US each will be made to submitters of entries judged "Commendation."

2p TUE. PM

Session 2pAB

Animal Bioacoustics: Methodology for Measurements of Auditory Evoked Potentials in Aquatic Mammals II

Mardi C. Hastings, Chair

Office of Naval Research, ONR 341, 800 North Quincy St., Arlington, VA 22217

Chair's Introduction—1:25

Contributed Papers

1:30

2pAB1. Effects of surveillance towed array sensor system (SURTASS) low frequency active sonar on fish. Arthur N. Popper, Michele B. Halvorsen, Diane Miller, Michael E. Smith, Jiakun Song, Lidia E. Wysocki (Dept. of Biol., Univ. of Maryland, College Park, MD 20742, apopper@umd.edu), Mardi C. Hastings (Office of Naval Res., Arlington, VA 22217), Andrew S. Kane (Univ. of Maryland, College Park, MD 20742), and Peter Stein (Sci. Solutions, Inc., Nashua, NH 03049)

We investigated the effects of exposure to Low Frequency Active (LFA) sonar on rainbow trout (a hearing non-specialist related to several endangered salmonids) and channel catfish (a hearing specialist), using an element of the standard SURTASS LFA source array. We measured hearing sensitivity using auditory brainstem response, effects on inner ear structure using scanning electron microscopy, effects on non-auditory tissues using general pathology and histopathology, and behavioral effects with video monitoring. Exposure to 193 dB re 1 microPa (rms received level) in the LFA frequency band for 324 seconds resulted in a TTS of 20 dB at 400 Hz in rainbow trout, with less TTS at 100 and 200 Hz. TTS in catfish ranged from 6 to 12 dB at frequencies from 200 to 1000 Hz. Both species recovered from hearing loss in several days. Inner ears sensory tissues appeared unaffected by acoustic exposure. Gross pathology indicated no damage to non-auditory tissues, including the swim bladder. Both species showed consistent startle responses at sound onsets and changed their position relative to the sound source during exposures. There was no fish death attributable to sound exposure even up to four days post-exposure. [Work supported by Chief of Naval Operations.]

1:45

2pAB2. The effects of stimulus parameters on the auditory brainstem response of zebrafish (*Danio rerio*). Maria Margaritis and Dennis Higgs (Dept. of Biol., Univ. of Windsor, Windsor, ON, Canada N9B 3P4, dhiggs@uwindsor.ca)

In mammals and birds it is well documented that stimulus parameters have significant effects on auditory processing. However in fish, the effect of different stimulus types remains unclear. Therefore it is difficult to directly compare piscine auditory responses obtained by different investigators, due to variability in testing methods. The current study uses zebrafish (*Danio rerio*) to evaluate the effects of stimulus type (800 Hz tone bursts versus broadband stimuli from 100–4000 Hz) and stimulus duration (1, 5, 10, or 20 ms) on the auditory brainstem response (ABR). Threshold was significantly lower ($p=0.04$) for broadband sounds than for tones, but there was no effect on either latency or amplitude. There was a direct and significant effect of duration on threshold ($p=0.001$), suggesting the response was due to the offset rather than the onset of the stimulus. Duration had no effect on either threshold or amplitude. The stimulus level had no significant effect on latency, however amplitude significantly increased as level increased ($p=0.001$). Thus, stimulus parameters can significantly affect ABRs. The current findings show the need for a consistent method of testing among investigators in order to reliably compare results. [Work supported by NSERC.]

2:00

2pAB3. Auditory evoked potential measurements in elasmobranchs. Brandon Casper and David Mann (Univ. of South Florida, College of Marine Sci., 140 Seventh Ave. South, St. Petersburg, FL 33701)

Auditory evoked potentials (AEP) were first used to examine hearing in elasmobranchs by Corwin and Bullock in the late 1970s and early 1980s, marking the first time AEPs had been measured in fishes. Results of these experiments identified the regions of the ear and brain in which sound is processed, though no actual hearing thresholds were measured. Those initial experiments provided the ground work for future AEP experiments to measure fish hearing abilities in a manner that is much faster and more convenient than classical conditioning. Data will be presented on recent experiments in which AEPs were used to measure the hearing thresholds of two species of elasmobranchs: the nurse shark, *Ginglymostoma cirratum*, and the yellow stingray, *Urobatis jamaicensis*. Audiograms were analyzed and compared to previously published audiograms obtained using classical conditioning with results indicating that hearing thresholds were similar for the two methods. These data suggest that AEP testing is a viable option when measuring hearing in elasmobranchs and can increase the speed in which future hearing measurements can be obtained.

2:15

2pAB4. Bottlenose dolphins audiogram dependence on azimuth: Evoked potential study. Vladimir Popov (Inst. of Ecology and Evolution, 33 Leninsky Prosp., 119071 Moscow, Russia)

ABR thresholds to tonal pips were measured in two bottlenose dolphins at different azimuthal positions of the sound source. The tested frequency range was from 8 to 128 kHz. Azimuth varied within a limit of 90 degree relative to the animals' longitudinal axis. This experimental paradigm allowed us to obtain ABR audiograms at different locations of the sound source. The zero-azimuth audiogram, at the sound source position in front of the animal, was of a standard appearance (minimum thresholds at frequencies of 38–90 kHz, steep threshold increase at higher frequencies, and shallower increase at lower frequencies). The audiograms at lateralized sound-source positions looked in a different manner. With the azimuth increase, high-frequency thresholds rose much higher than low-frequency ones, so at azimuths of 60–90, the threshold versus frequency function was almost monotonous: the lowest threshold was observed at the lowest frequency (8 kHz) and the highest threshold at the highest frequency (128 kHz). With monaural ABR recording, audiograms contralateral to the sound source featured higher thresholds and steeper threshold increase with frequency as compared to the ipsilateral ones. [Work supported by the Russian Foundation for Basic Research.]

2:30

2pAB5. Audiogram variability in normal bottlenose dolphins.

Mikhail B. Tarakanov, Mikhail G. Pletenko, Vladimir V. Popov, and Alexander Ya. Supin (Inst. of Ecology and Evolution, 33 Leninsky Prospect, 119071 Moscow, Russia)

Audiograms have been obtained in about a dozen of odontocete species, but mostly in one or two individuals each. However, some inter-individual difference in hearing sensitivity is inevitable. Therefore, a representative number of animals should be investigated to get a normal audiogram standard. In the present study, an attempt has been made to estimate the audiogram scatter among normal bottlenose dolphins. Measurements were made in dolphins captured in wild and kept in captivity 3 to 5 months, using auditory evoked potential technique (envelope following response) to measure hearing thresholds in far acoustic field. Seven subjects, 5 males and 2 females, provisionally from 3 to 15 years old, were investigated during the summer season of 2004. Hearing thresholds were measured at frequencies from 8 to 152 kHz with quarter-octave steps. All the subjects had qualitatively similar audiograms. The best sensitivity was from 38.9 dB re 1 uPa (at 32 kHz) to 51.9 dB (at 16 kHz), with a minimum of the averaged audiogram of 47.1 dB at 45 kHz. High-frequency cut-off was 152 kHz at a level of 40 dB above the lowest threshold. Standard deviation of threshold was from 4.1 to 10.3 dB. [Work supported by RFBR and Russian President Grants.]

2:45

2pAB6. Auditory evoked potential (AEP) measurements in stranded rough-toothed dolphins (*Steno bredanensis*).

Mandy L. H. Cook (USF College of Marine Sci., 140 Seventh Ave. South, Saint Petersburg, FL 33701-5016, mhill@marine.usf.edu), Charles A. Manire (Mote Marine Lab., Sarasota, FL 34236), and David A. Mann (USF College of Marine Sci., Saint Petersburg, FL 33701-5016)

Thirty-six rough-toothed dolphins (*Steno bredanensis*) live-stranded on Hutchinson Island, FL on August 6, 2004. Seven animals were transported to Mote Marine Laboratory for rehabilitation. Two auditory evoked potential (AEP) measurements were performed on each of five of these dolphins in air using a jawphone to present acoustic stimuli. Modulation

rate transfer functions (MRTFs) were measured to establish how well the auditory system follows the temporal envelope of acoustic stimuli. A 40 kHz stimulus carrier was amplitude modulated (AM) with varying rates ranging from 200 Hz to 1800 Hz, in 200 Hz steps. The best AM-rate from the first dolphin tested was 1500 Hz. This AM rate was used in subsequent AEP measurements to determine evoked-potential hearing thresholds between 5000 and 80 000 Hz. These findings show that rough-toothed dolphins can detect sounds between 5 and 80 kHz, and are most likely capable of detecting frequencies much higher than 80 kHz. MRTF data suggest that rough-toothed dolphins have a high temporal resolution, similar to that of other cetaceans.

3:00

2pAB7. Stimulus-response characteristics of a harbor porpoise during active echolocation and passive hearing studied with auditory brain stem recordings (ABR).

Kristian Beedholm and Lee A. Miller (Inst. of Biol., Univ. of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark, lee@biology.sdu.dk)

We evaluated the stimulus-response characteristics of hearing by a harbor porpoise using narrow band pulses resembling the animal's own echolocation pulse (130 kHz, 100 us), but shifted in frequency (80, 100, 125, 160 kHz). Our animal was trained to accept two suction cup electrodes and to station 1 m below the water surface. Stimuli could be presented either as simulated echoes at a fixed delay (5 ms) relative to the animal's echolocation signals, or at a constant rate chosen by the experimenter. Stimulus levels were varied between 90 and 150 dB re 1uPa and the ABR responses were averaged (16 or more responses) at each level. The relationship between input level (in dB) and ABR amplitude was reasonably linear for simulated echo and constant rate experiments. Regression lines were calculated to determine the level at which the response met the noise, defining the ABR threshold. There was little difference in the ABR threshold (100 to 110 dB re 1uPa) for the four frequencies. The rate of growth of the ABR response with increasing stimulus level was steepest at 125 kHz, which could well reflect a relatively denser population of neurons tuned to this frequency area. [Work done at Fjord & Baelt, Kerteminde, Denmark, and supported by ONR.]

TUESDAY AFTERNOON, 17 MAY 2005

PLAZA A, 1:30 TO 5:00 P.M.

Session 2pAO

Acoustical Oceanography and Underwater Acoustics: Geoacoustic Inversion Techniques

Michael J. Buckingham, Chair

Scripps Inst. of Oceanography, 8820 Shellback Way, La Jolla, CA 92093-0238

Contributed Papers

1:30

2pAO1. Seafloor classification using artificial neural network architecture from central western continental shelf of India.

Vasudev Mahale, Bishwajit Chakraborty, Gajanan S. Navelkar, and R. G. Prabhu Desai (Natl. Inst. of Oceanogr., Dona Paula, Goa: 403 004, India, bishwajit@darya.nio.org)

Seafloor classification studies are carried out at the central western continental shelf of India employing two frequency normal incidence single beam echo-sounder backscatter data. Echo waveform data from different seafloor sediment areas are utilized for present study. Three arti-

ficial neural network (ANN) architectures, e.g., Self-Organization Feature Maps (SOFM), Multi-Layer Perceptron (MLP), and Learning Vector Quantization (LVQ) are applied for seafloor classifications. In case of MLP, features are extracted from the received echo signal, on the basis of which, classification is carried out. In the case of the SOFM, a simple moving average echo waveform pre-processing technique is found to yield excellent classification results. Finally, LVQ, which is known as ANN of hybrid architecture is found to be the efficient seafloor classifier especially from the point of view of the real-time application. The simultaneously acquired sediment sample, multi-beam bathymetry and side scan sonar and echo waveform based seafloor classifications results are indicative of the

2p TUE. PM

depositional (inner shelf), non-depositional or erosion (outer shelf) environment and combination of both in the transition zone. [Work supported by DIT.]

1:45

2pAO2. Geoacoustic parameters inversion by ambient noise in South China Sea experiment of ASIAEX. Ruey-Chang Wei, Yao-Hsien Kuo, Cheng-Hung Chen (Natl. Sun Yat-Sen Univ., No. 70, Lien-Hai Rd., Kaohsiung, Taiwan 804), and Hsiang-Chih Chan (Natl. Taiwan Univ., Taipei, 106, Taiwan)

Knowing bottom geoacoustic parameters is of great importance for using sonar systems effectively in shallow water. This study is based upon the vertical line array (VLA) in South China Sea experiment of the Asian Seas International Acoustic Experiment (ASIAEX), dated from May 3, 2001 to May 16, 2001, in which the ambient noise was also measured in shallow water. Beamforming is used to calculate the vertical directionality of ambient noise, but limited by the array element spacing, the major analysis frequency range is 100 to 400 Hz. Comparison between forward modeling (normal mode code) and experimental results show the energy ratio of up-looking and down-looking beams, which can be referred to the energy reflection coefficient of the bottom-potentially a function of angle and frequency. From the reflection coefficient, critical parameters of the sediments such as sound speed, density, and attenuation coefficient can be obtained. Finally, this result is verified by on-site chirp sonar surveys during the experiment, and also by other inversion investigation in this area using other sound source.

2:00

2pAO3. Single beam echosounding: Considerations of depth and seabed slope. Ben R. Biffard, N. Ross Chapman, Steve F. Bloomer (School of Earth and Ocean Sci., Univ. of Victoria, P.O. Box 3055, Victoria, BC, Canada V8W 3P6), and Jon M. Preston (Quester Tangent Corp., Sidney, BC, Canada V8L 5Y8)

The topic of depth compensation of single beam echo time series for seabed classification is fairly well studied. The effect of seabed slope has not been publicized. In applications for seabed classification, seabed slope is observed to be associated with classification inaccuracy and failure. In cases of higher slope, single beam bathymetry also becomes inaccurate. A survey of 2 fjords with extreme variation in slope is presented as a representative example and testing bed for investigating slope. The direct effect on seabed echoes is investigated and explained in reference to a simple model of beam-echo geometry. Survey bathymetry is compensated for slope; inaccuracies of up to 5% of depth are corrected, however bottom picking accuracy is diminished in areas of slope and cannot be improved. Surveys of a gas hydrate site and a river will also be presented as applications of these ideas. Early results from BORIS model studies of methods for compensation of slope and depth may also be presented.

2:15

2pAO4. Estimating the uncertainty of geoacoustic parameters of a range-dependent environment. Mark Fallat (Dept. of Earth and Ocean Sci., Univ. of BC, Canada), Stan Dosso (Univ. of Victoria, Canada), and Peter Nielsen (NATO Undersea Res. Ctr.)

In this study, data from a range-dependent environment are inverted to obtain estimates of the geoacoustic properties and their uncertainties. The technique consists of combining the results of a series of range-independent inversions to produce a model of the range-dependent environment. A rigorous uncertainty analysis provides a way of discerning whether variations in the inversion results are due to range-dependent features of the environment or simply due to uncertainty/variability in the results. Broadband acoustic data from a track off the island of Sicily were analyzed. Ground-truth information in the form of core measurements and a high-resolution seismic profile were also collected and used for verification of the results. The method of fast Gibbs sampling (FGS) was used to estimate the uncertainties of the geoacoustic properties. FGS is based on a Bayesian approach to inversion which samples the posterior probability

distribution to estimate marginal probability distributions and parameter correlations. Marginal probability distributions were computed at various points along the track and compared to the ground-truth information. Overall, the analysis showed that the dominant range-dependent features of the environment could be estimated.

2:30

2pAO5. A method for obtaining the geoacoustic properties of marine sediments using sound generated by aircraft. Eric M. Giddens and Michael J. Buckingham (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0238)

A low-frequency geoacoustic inversion method has been developed for an isovelocity waveguide using light-aircraft as a source of sound. The high Doppler shift of the acoustic field facilitates the inversion, allowing directional information to be gathered by a single, omnidirectional receiver. The inversion method has been applied to simulations and to field experiments conducted in shallow water (14.4 m depth) off the coast of La Jolla. The inversion results will be discussed along with an error analysis using the Cramer-Rao Lower Bounds. [Work supported by ONR and the ARCS Foundation.]

2:45

2pAO6. Uncertainty bounds for geoacoustic inversions due to sound speed variations in water column. Ying-Tsong Lin and James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

The uncertainty bounds for geoacoustic inversion results due to sound speed variability in the water column perturbed by internal waves are assessed. Our previous work showed that if the sound speed variations are not taken account into the water column model, they will affect the accuracy of geoacoustic inversions. In this paper, we seek to quantify how well the error bounds could be given under certain knowledge of the sound speed variations. In doing so, the equivalent transform method derived from the linear perturbation theory plays a role in error transferral from water column to bottom. The water column variability is decomposed into ocean dynamic modes or as empirical orthogonal function (EOF) modes. Obtaining the ocean dynamic modes requires having the buoyancy frequency profile, whereas obtaining the EOF modes requires having vertical time series data. In general, it is difficult to assemble full statistics of the sound speed variations, so we examine the uncertainty bounds of the inversion results in the case where only partial statistics can be observed.

3:00

2pAO7. Geoacoustic inversion of broadband data by matched beam processing. N. Ross Chapman and Yongmin Jiang (School of Earth and Ocean Sci., Univ. of Victoria, P.O. Box 3055, Victoria, BC, Canada V8W 3P6)

This paper describes results of geoacoustic inversion using broadband signals from an experiment carried out at a site near the South Florida Ocean Measurement Centre in the Florida Straits. M-sequence coded pulse trains at different center carrier frequencies from 100 to 3200 Hz were transmitted in the acoustic experiment, and received by a tri-axial array at a distance around 10 km. Geoacoustic inversion was carried out to determine the possibility of inverting the environmental parameters from this long-range propagation experiment. The received signal at 100 Hz consisted of a dominant water column signal and a secondary arrival delayed by 0.4 s. The secondary signal was spatially filtered by beamforming the array data, and the beam data were inverted by matched beam processing combined with an adaptive simplex simulated annealing algorithm. The estimated values of compressional wave speed and density were in good agreement with ground truth values from sediment cores. The inverted shear wave speed appears to be a sensitive parameter and consistent with compressional wave speed. Range and water depth were also included as inversion parameters, and the inversion results were close to the known values within small uncertainties. [Work supported by ONR.]

3:15

2pAO8. Matched-field geoaoustic inversion with correlated data errors. Stan E. Dosso (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC, Canada)

Many approaches to geoaoustic inversion are based implicitly on the assumption that data errors (including measurement and theory errors) are uncorrelated. However, this assumption is often not valid, and can lead to inefficient geoaoustic parameter estimates and an under-estimation of parameter uncertainties. This paper examines the effects of data error (residual) correlations in matched-field geoaoustic inversion. Statistical tests for error correlations in both space and frequency are considered. An inversion approach is developed based on an iterative, non-parametric method of estimating the full error covariance matrix (assuming ergodic residuals), and including this covariance in the likelihood function of a Bayesian formulation. The inversion approach is illustrated for broadband, vertical-array data from the Mediterranean Sea.

3:30–3:45 Break

3:45

2pAO9. Bounding the error on bottom estimation for multi-angle swath bathymetry sonar. Geoff K. Mullins and John S. Bird (School of Eng. Sci., Simon Fraser Univ., 8888 University Dr., Burnaby, BC, Canada V5A 1S6)

With the recent introduction of multi-angle swath bathymetry (MASB) sonar to the commercial marketplace (e.g., Benthos Inc., C3D sonar, 2004), additions must be made to the current sonar lexicon. The correct interpretation of measurements made with MASB sonar, which uses filled transducer arrays to compute angle-of-arrival information (AOA) from backscattered signal, is essential not only for mapping, but for applications such as statistical bottom classification. In this paper it is shown that aside from uncorrelated channel to channel noise, there exists a tradeoff between effects that govern the error bounds on bottom estimation for surfaces having shallow grazing angle and surfaces distributed along a radial arc centered at the transducer. In the first case, as the bottom aligns with the radial direction to the receiver, footprint shift and shallow grazing angle effects dominate the uncertainty in physical bottom position (surface aligns along a single AOA). Alternatively, if signal from a radial arc arrives, a single AOA is usually estimated (not necessarily at the average location of the surface). Through theoretical treatment, simulation, and field measurements, the aforementioned factors affecting MASB bottom mapping are examined. [Work supported by NSERC.]

4:00

2pAO10. Posterior probability distributions for Biot parameters from experimental data. Buye Xu, Traci Neilsen (Dept. of Phys. and Astron., Brigham Young Univ., N203 ESC, Provo, UT 84602, bxu@email.byu.edu), and Marcia Isakson (Univ. of Texas, Austin, TX 78712)

Reflectivity measurements can be used in inversion techniques to obtain estimates of the physical characteristics of seabed sediments. In addition, computation of posterior probability distributions (PPDs) [S. E. Dosso, *J. Acoust. Soc. Am.* **111**, 129–142 (2002)] provide an uncertainty of each parameter estimate obtained by a simulated annealing inversion. In this work, PPDs are calculated for experimental reflection loss data to evaluate the sensitivities associated with the parameters in the Biot-Stoll poro-elastic model for sediments. The data comes from a lab tank experiment in which reflection loss from a sandy sediment was measured as a function of grazing angle for frequencies between 75 kHz and 150 kHz. The sampling for the PPDs is based on the least-squares error between simulated reflection coefficient data and values modeled with the reflection module of OASES. [H. Schmidt, “OASES Version 2.1 User Guide

and Reference Manual,” Department of Ocean Engineering, Massachusetts Institute of Technology, 1997.] The PPDs resulting from the lab data are compared with those obtained for simulated data, and the effects of experimental error on the results are explored.

4:15

2pAO11. Measurement of the bulk density gradient in the transition layer of a sand/water interface using a diver deployed X-Ray Attenuation Measurement System (XRAM). Todd Hay and Marcia Isakson (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713)

The reflection coefficient is an important parameter for modeling acoustic applications in littoral environments. However, a density gradient in the transition layer of the interface can change the value of the reflection coefficient by 15 dB for high frequencies. Therefore, an accurate measurement of the bulk density gradient is crucial to the interpretation of reflection coefficient data. Previously, bulk density gradients have been determined by coring which disrupts grain microstructure or by electrical methods which do not measure the density directly. For this study, a novel means of directly determining the density gradient in situ using x-ray attenuation was developed. The system is small, lightweight, and diver deployed allowing a greater flexibility in density measurements. Calibration measurements of materials with varying densities as well as measurements of the transition layer of a sand/water interface will be presented. [Work supported by ONR, Ocean Acoustics.]

4:30

2pAO12. Geoaoustic inversion for single paths of Haro Strait data. A. Tolstoy (A. Tolstoy Sci., 1350 Beverly Rd., Ste. 115 PMB 293, McLean, VA 22101)

This presentation will discuss efforts to date for the geoaoustic inversion of selected, single, Haro Strait Source-Receiver (SR) paths. The parameters emphasized will include water depth, linear sound-speed profiles, and the range between SR. The path will assume 3 segments: one segment for the region at S, one for the region at R, and one for an average region in between. The inversions will operate using the SUB-RIGS method which is based on multiple frequencies where higher frequencies are most sensitive to surficial geoaoustic properties while the lower frequencies are most sensitive to the deeper properties. Some paths have already been successfully inverted by earlier research efforts (Chapman *et al.*). This paper will demonstrate efforts on some of those paths as well as on other, more problematic SR paths.

4:45

2pAO13. Classification of acoustic signals using the statistics of the 1-D wavelet transform coefficients. Michael I. Taroudakis (Dept. of Mathematics, Univ. of Crete and FORTH/IACM, P.O. Box 1527, 711 10 Heraklion, Greece) and George Tzagkarakis (Univ. of Crete and FORTH/ICS, Heraklion, Greece)

The representation of an acoustic signal in terms of the wavelet sub-band coefficients is studied using their statistical features. This type of representation is suggested as an alternative tool for tomographic or geoaoustic inversions. Using a set of representative shallow water environments, the variation of the statistical behavior of the sub-band coefficients is associated with corresponding variations of the environmental parameters. Low frequency signals suitable for ocean acoustic tomography are simulated. It is shown that the statistics of the sub-band coefficients are best described using non-Gaussian heavy-tailed distributions such as those of the alpha-stable family. The variations of the distribution parameters are measured using special statistical similarity functions such as the Kullback-Leibler divergence which has been adopted in the present study. The distances between the statistical parameters of a given and reference signals determined by these functions, can be used for subsequent non-linear inversions aiming at the recovery of the environmental parameters.

2p TUE. PM

Session 2pBB

Biomedical Ultrasound/Bioresponse to Vibrational and Physical Acoustics: Ultrasound Imaging

Jeffrey A. Ketterling, Cochair

Riverside Research Inst., 156 William St., New York, NY 10038-2609

James C. Lacefield, Cochair

Dept. of Electrical and Computer Engineering, Univ. of Western Ontario, London, ON N6A 5B9, Canada

Chair's Introduction—1:45

Contributed Papers

1:45

2pBB1. Computational synthesis of ultrasound breast images from a three-dimensional anatomical model. Yi-Ting Shen and James C. Lacefield (Univ. of Western Ontario and Robarts Res. Inst., London, ON, Canada)

A three-dimensional breast anatomy model has been implemented using spline surfaces and fractal structures to represent the architecture of the lactiferous ducts, mammary fat lobules, skin, and supporting connective tissues. The model randomly varies user-specified structural parameters to provide an unlimited number of realizations of the gross anatomy. Cross-sectional views extracted by slicing through a realization of the model are input to a two-dimensional k -space (i.e., spatial frequency domain) ultrasound propagation simulator. The k -space simulator iterates pressure and particle velocity fields in 30-ns steps to compute scattering from the structures defined by the anatomical model and small random variations in compressibility that are added to generate speckle. A synthetic aperture method is employed to simulate B -mode imaging with a 5 MHz, 192-element linear array operated using multiple transmit focal zones and dynamic receive focusing. Simulated images of random-scattering phantoms possess approximately Rayleigh speckle statistics. The anatomical model is expected to yield images with speckle statistics comparable to clinical breast images. The long-term objectives of these simulations are to investigate sources of focus aberration in ultrasound breast imaging and the impact of aberration on cancer detection. [Work supported by an NSERC Discovery Grant.]

2:00

2pBB2. Tissue-type imaging (TTI) based on ultrasonic spectral and clinical parameters for detecting, evaluating, and managing prostate cancer. Ernest J. Feleppa, Jeffrey A. Ketterling, Shreedevi Dasgupta, Andrew Kalisz, Sarayu Ramachandran (Biomed. Eng. Labs., Riverside Res. Inst., 156 William St., New York, NY 10038, feleppa@rrinyc.org), and Christopher R. Porter (Virginia Mason Medical Ctr., Seattle, WA 98101)

This study seeks to develop more-sensitive and -specific ultrasonic methods of imaging cancerous prostate tissue and thereby to improve means of guiding biopsies and planning, targeting, and monitoring treatment. Ultrasonic radio-frequency, echo-signal data, and clinical variables, e.g., PSA, voiding function, etc., during biopsy examinations were acquired. Spectra of the radio-frequency signals were computed in each biopsied region, and used to train neural networks; biopsy results served as the gold standard. A lookup table gave scores for cancer likelihood on a pixel-by-pixel basis from locally computed spectral-parameter and global clinical-parameter values. ROC curves used leave-one-patient- and leave-one-biopsy-out approaches to minimize classification bias. Resulting ROC-curve areas were 0.80 ± 0.03 for neural-networks versus 0.66 ± 0.03 for conventional classification. TTIs generated from data acquired pre-surgically showed tumors that were unrecognized in conventional images

and during surgery. 3-D renderings of prostatectomy histology and TTIs showed encouraging correlations, which shows promise for improving the detection and management of prostate cancer, e.g., for biopsy guidance, planning dose-escalation and tissue-sparing options for radiation or cryotherapy, and assessing the effects of treatment. Combining MRS parameters with US spectral parameters appears capable of further improving prostate-cancer imaging. [Work supported by NIH.]

2:15

2pBB3. Bioheat Transfer Model (BHTE) based temperature estimation technique for high intensity focused ultrasound therapy monitoring. Peter J. Kaczowski and Ajay Anand (Appl. Phys. Lab, Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

The spatial distribution of tissue temperature is an essential indicator of thermal therapy progress, treatment safety and efficacy. Here, it is shown through in vitro experiments that temperature rise can be accurately measured during therapy delivery and post-treatment cool down phases using RF backscatter data collected with a modified ultrasound scanner. RF data is acquired prior to, during, and after HIFU exposures, in tissue mimicking phantoms and excised animal tissue. Through two calibration experiments, initial estimates for key BHTE parameters (local thermal diffusivity, and magnitude of the HIFU heat source) and the temperature dependence of ultrasonic travel time are obtained prior to therapy. Tissue heterogeneity over a larger targeted region of interest is modeled as a change in the magnitude of the focal heat source. During therapy this magnitude is updated using an iterative optimization technique that minimizes the difference between predicted and measured travel time values. Temperature and thermal dose maps are generated throughout therapy delivery and post-treatment cooling periods. The ultrasound derived estimates are validated against independent thermocouple measurements close to but not at the HIFU focus. This model-based technique permits noninvasive temperature estimation throughout the entire therapeutic range, and is thus a departure from previously reported techniques.

2:30

2pBB4. Comparison of three models at high frequency for ultrasound tissue scattering. Michael Oelze, William O'Brien, Jr. (Dept. of Elec. and Comput. Eng., Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801), and James Zachary (Univ. of Illinois at Urbana-Champaign)

A mammary carcinoma was grown in mice and imaged with an ultrasound transducer operating with a center frequency of 65 MHz. Quantitative ultrasound (QUS) analysis was used to characterize the tumors using the bandwidth of 30 to 85 MHz. Three models (Gaussian scatterer, fluid-filled sphere, and a new cell scatterer) for scattering were examined and scatterer property estimates were compared to real tissue morphology as seen from optical microscope images of the tumors. The Gaussian scattering model did not fit the data well compared to the fluid-filled sphere and new cell scatterer models. The fluid-filled sphere model fit the measure-

ments better than any other model but did not yield scatterer property estimates that resembled underlying structure. Using the fluid-filled sphere model, the average estimated scatterer diameter was $25.5 \pm 0.14 \mu\text{m}$. A new cell scatterer model was developed, which was based on scattering from a cell by incorporating the effects of the cytoskeleton and nucleus. The new cell scatterer model yielded estimates that appeared to reflect underlying structure more accurately. Using the new cell scatterer model, the average estimated nuclear diameter was $15.6 \pm 2.2 \mu\text{m}$ compared with $13.2 \mu\text{m}$ as measured from optical microscope images.

2:45

2pBB5. Importance of local attenuation on the estimation of scatterer size from ultrasound backscattered waveforms. Timothy A. Bigelow and William D. O'Brien, Jr. (Univ. of Illinois, 405 N. Mathews Ave., Urbana, IL 61801, bigelow@uiuc.edu)

Estimating the characteristic size of the tissue microstructure from a statistical analysis of the backscattered power spectrum could improve the diagnostic capability of medical ultrasound. Previously, size estimates were obtained for focused sources after compensating for source focusing, the frequency-dependent attenuation along the propagation path (total attenuation), and the frequency-dependent attenuation in the region of interest (local attenuation) given the attenuation values. In this study, the importance of the local attenuation on the scatterer size estimate was determined using computer simulations. The simulations used Gaussian impedance distributions with an effective radius of $25 \mu\text{m}$ randomly positioned in a homogeneous half-space (attenuation from 0.05 to 1 dB/cm-MHz) sonified by a spherically focused source ($f/1$ to $f/4$). The total attenuation and focusing were assumed known. The scatterer size was estimated assuming that the local attenuation was zero, the local attenuation was the mean value of the true attenuation over the frequency range used to obtain the estimates, and the local attenuation was 0.5 dB/cm-MHz. These three size estimates were then compared to estimates obtained using the true local attenuation value as well as estimates obtained when both focusing and local attenuation were not compensated.

3:00–3:15 Break

3:15

2pBB6. Non-invasive measurement of *in situ* thermal diffusivity and local heat source using backscattered ultrasound for thermal therapy planning and monitoring. Ajay Anand and Peter Kaczkowski (Univ. of Washington, 1013 40th St NE, Seattle, WA 98105)

Bioheat transfer equation (BHTE) estimates of applied dose during HIFU therapy typically use *prior* knowledge or assume standard values for tissue properties that determine thermal diffusivity (K) and heat source (Q). We have developed a novel signal-processing based technique to *noninvasively* estimate these parameters *in situ* based on analysis of raw backscattered RF data from two localized HIFU exposures, one at sub-ablative intensities to determine K , and another at therapeutic intensities to obtain Q . Both exposures are performed prior to therapy. To estimate K , a short HIFU exposure is applied resulting in a temperature rise of less than 15°C . The Gaussian radius of the temperature induced apparent strain profile during cool down is estimated and fit to an approximate analytical heat diffusion expression to obtain K . Independent estimates of K derived using the transient hotwire technique validated the ultrasonic measurements. To estimate Q , an audio range hydrophone was acoustically coupled to the sample to detect the onset of boiling. The time required to bring the sample to boiling was used to estimate Q by iteration of a numerical BHTE model. These results are validated against *in situ* measured values using thermocouples and linear acoustic calculations.

2pBB7. Validation of high-frequency ultrasound measurements of tissue layer thickness. Qiang Qiu, Joy Dunmore-Buyze, Derek R. Boughner, and James C. Lacefield (Univ. of Western Ontario, Robarts Res. Inst., and London Health Sci. Ctr., London, ON, Canada)

High-frequency ultrasound imaging enables nondestructive measurement of layer thickness in tissue specimens. These measurements are valuable for mechanical testing of soft biomaterials. This paper demonstrates a method for assessing the accuracy of high-resolution ultrasonic thickness estimates. Three-dimensional images of six porcine aortic valve cusps were acquired *in vitro* using a 40 MHz ultrasound system with $40 \times 80 \times 80 \mu\text{m}^3$ spatial resolution. The cusps were then frozen in liquid nitrogen, sectioned into $10\text{-}\mu\text{m}$ slices, and micrographs of one slice from each specimen were acquired at $4\times$ magnification. The two-dimensional micrographs were registered to the three-dimensional ultrasound images using a cross-correlation method. The boundaries of the fibrosa, spongiosa, and ventricularis layers were segmented in both sets of images using an active contour model. The average thicknesses of the tissue layers in the registered images were estimated and the absolute differences of the optical and ultrasonic estimates were computed. The absolute differences were $55.8 \pm 22.6 \mu\text{m}$ (mean \pm standard deviation), $23.5 \pm 14.3 \mu\text{m}$, and $22.7 \pm 17.2 \mu\text{m}$ for the fibrosa, spongiosa, and ventricularis, respectively. The measurement differences are comparable to the axial resolution of the ultrasound system and are not significant as determined by *t*-tests ($p > 0.30$ for each layer).

3:45

2pBB8. Automatic 3D acoustic tissue models from histologic tissue sections and application to *ex vivo* tissue characterization. Jonathan Mamou, Michael L. Oelze, William D. O'Brien, Jr. (Dept. of Elec. and Comput. Eng., Univ. of Illinois at Urbana-Champaign, 1406 W. Green St., Urbana, IL 61801, mamou@uiuc.edu), and James F. Zachary (Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801)

Three-dimensional acoustic tissue models (3DATMs) can be used as computational tools for ultrasonic imaging algorithm development and analysis. 3DATMs are automatically constructed from digitized light microscope images of consecutive H&E-stained histologic tissue sections. Construction necessitated contrast equalization, registration, and interpolation of missing sections. The registered (with interpolated) sections yield a 3D histologic volume (3DHV). Acoustic properties are then assigned to each tissue constituent of the 3DHV to obtain the 3DATM. A tissue characterization technique was developed to obtain scatterer parameter estimates (size and acoustic concentration) from a 3D impedance map (3DZM) deduced from a 3DHV by assigning acoustic impedance values. 3DZMs were constructed for a rat fibroadenoma (FA), a mouse mammary tumor (MMT) and a mouse sarcoma (EHS). From these 3 3DZMs estimates, effective scatterer diameters of $91 \mu\text{m}$, $31.5 \mu\text{m}$, and $34.5 \mu\text{m}$, respectively, were determined. Independent ultrasonic measurements yielded average scatterer diameters of $105 \mu\text{m}$, $30 \mu\text{m}$, and $33 \mu\text{m}$, respectively. The 3DZM estimation scheme showed results similar to those obtained by the ultrasonic measurements. 3DATMs may therefore be a useful tool for quantifying ultrasonic tissue properties. [Work supported by the University of Illinois Research Board.]

4:00

2pBB9. Automated set of test simulations for non-ideal medical ultrasound transducers using Monte-Carlo method. Hotaik Lee, Nadine B. Smith (Grad. Program in Acoust., The Penn State Univ., University Park, PA 16802), and Terry A. Kling (Sound Technol. Inc., State College, PA 16803)

Using Field II program and Monte-Carlo method, we developed new design tools for medical ultrasound transducers with the automated set of imaging simulations. This simulation environment is used to create and assess the parametric specification of design factors affecting the quality of medical ultrasound imaging. In order to obtain accurate and realistic results, non-ideal transducers whose transfer functions vary either across the transducer or within the transducer are considered. These variations

include, but are not limited to: center frequency, bandwidth, sensitivity, ringdown, angular response, time-of-flight, and lateral focus. By applying random numbers within the tolerance range to the variations in input parameters, the automated set of simulations is performed. First, critical input parameters for components of the transfer function are identified by sensitivity analysis. Next, the statistical range of parameter values that

yield a transducer model with a certain performance level is determined and the limit of variations in each factor for acceptable degradation of images is set. Finally, the creation of many “what if” cases to predict yield and statistical performance of a transducer and the imaging simulation are performed based on Monte-Carlo method. [Work supported by Sound Technology Inc.]

TUESDAY AFTERNOON, 17 MAY 2005

GEORGIA A, 1:00 TO 4:05 P.M.

Session 2pEA

Engineering Acoustics, Acoustical Oceanography and Underwater Acoustics: Underwater Acoustic Sensor Technologies

Dehua Huang, Cochair

Naval Undersea Warfare Center, 1176 Howell St., Newport, RI 02841-1708

Thomas R. Howarth, Cochair

NAVSEA Newport, 1176 Howell St., Newport, RI 02841

Chair's Introduction—1:00

Invited Papers

1:05

2pEA1. Injection-molded 1-3 piezocomposite sensor development: The last ten years. Kim C. Benjamin (Naval Sea Systems Command Div. Newport, 1176 Howell St., Newport, RI 02841)

The past ten years have seen several interesting demonstrations of 1-3 piezocomposite when used as the active component in sonar sensors. Initially considered receive only by most in the field, piezocomposite has evolved into a proven broadband transducer material with both receive and transmit capability. From large aperture single element calibration transducers, to parametric mode projectors, the material has surprised many experts with its power handling capability. Its polymer constituent provides an amazing degree of versatility by allowing the thermoforming and shaping of transducer substrates for packaging into today's undersea vehicles. This talk will review the last ten years of piezocomposite transducer and array development focusing on both the materials transmit behavior and fabrication benefits for future sonar applications. [Work supported by the U.S. Navy.]

1:30

2pEA2. Engineering applications of limited diffraction beams. Jian-yu Lu (Ultrasound Lab, Dept. of Bioengineering, The Univ. of Toledo, Toledo, OH 43606, jilu@eng.utoledo.edu)

Limited diffraction beams (LDBs) are non-divergence and non-dispersive solutions to isotropic-homogeneous wave equations. These beams have a common characteristic of X-shaped branches and thus are also called X waves. Because of their highly directional propagation property, they have potential applications in medicine, underwater acoustics, and nonlinear optics (Charles Day, *Phys. Today*, October, 2004, pp. 25–26). In this talk, an overview of the development of LDBs will be given. This includes the conversion of any existing solutions to homogeneous or non-homogeneous wave equations to LDB solutions using Lorentz-related transformation, and a discussion of the orthogonal properties of the X wave transformation pair in representing any physically realizable waves. Experiments on applications of LDBs using our newly developed general-purpose ultrasound system will also be reported. (The system is capable of 128-channel simultaneous ultrasound data acquisitions at 12-bit/40 MHz rate and can hold real-time radio-frequency (RF) data up to 64 GB in one acquisition operation. Arbitrary ultrasound waveforms can be simultaneously produced by 128 12-bit/40 MHz D/A converters and then be linearly amplified to drive 75-Ohm loads or an array transducer at about ± 150 V.)

1:55

2pEA3. Fiber optic acoustic sensor technology. James Cole, Clay Kirkendall, and Anthony Dandridge (Naval Res. Lab, 4555 Overlook Ave., SW, Washington, DC 20375)

Fiber optic sensor technology has been under development for over 25 years, recently a major milestone has been reached- the introduction of the Fiber Optic Wide Aperture Array on the first Virginia class submarine. This paper will review the development of this technology, outlining the principles of operation and the technological developments that led to fiber optic interferometric sensors becoming viable for production in an advanced sonar system. The Fiber Optic Wide Aperture array is a large channel count planar array mounted on the side of the submarine, but fiber sensor technology is also being developed for both towed arrays (as a replacement for the Navy's thin-line towed arrays) and for bottom mounted acoustic arrays for a number of Navy applications. This

paper will describe the development of the fiber optic hydrophones for these applications as well as the optical interrogation techniques. One of the key features of fiber optic sensor technology is the ability to passively multiplex many hydrophone channels per fiber, several multiplexing techniques will also be described. Some of the issues and challenges of this technology, such as coherent noise for large channel count systems will also be briefly discussed.

Contributed Papers

2:20

2pEA4. Twenty years of barrel-stave flextensional transducer technology in Canada. Dennis F. Jones (Defence R&D Canada—Atlantic, P.O. Box 1012, Dartmouth, NS, Canada B2Y 3Z7, dennis.jones@drdc-rddc.gc.ca)

The barrel-stave flextensional transducer, a compact underwater sound source, was conceived at DRDC Atlantic in 1986 [G. W. McMahon and D. F. Jones, U.S. Patent No. 4,922,470 (1 May 1990); Canadian Patent No. 1,285,646 (2 July 1991)]. Over the years, five barrel-stave designs belonging to three flextensional classes were built and tested at DRDC Atlantic. Three Class I transducers with operating frequencies ranging from 800 to 1600 Hz were integrated into submarine communications buoys, low frequency active horizontal projector arrays, and a broadband sonar towbody. A high-power Class II and broadband (1–7 kHz) Class III transducer were deployed under the ice in the Lincoln Sea for research related to rapidly deployable surveillance systems. These barrel-stave flextensional transducers have also supported a variety of marine mammal studies including vocal mimicry in long-finned pilot whales, coda dialects in sperm whales, and the R&D of acoustic detection and tracking systems for endangered northern right whales. In August 2004 a barrel-stave transducer was used to lure a trapped juvenile humpback whale to the sluice gates of a tidal generating station on the Annapolis River in Nova Scotia by transmitting humpback whale calls underwater. The acoustic performance parameters for all 5 transducers will be presented.

2:35–2:50 Break

2:50

2pEA5. Single crystal cylinder transducers for sonar applications. Harold Robinson, Gerald Stevens, Martin Buffman (NUWC Div. Newport, 1176 Howell St., Newport, RI 02841), and James Powers (EDO Corp., Salt Lake City, UT 84115)

A segmented cylinder transducer constructed of single crystal lead magnesium niobate-lead titanate (PMN-PT) has been under development at NUWC and EDO Corporation for several years. The purpose of this development was to provide an extremely compact, high power broadband source. By virtue of their extraordinary material properties, ferroelectric single crystals are the ideal transduction material for developing such compact broadband systems. This presentation shall review the evolution of the transducer design as well as present the results of a successful in-water test conducted at NUWC in October of 2003. It shall be shown that design changes intended to eliminate spurious modes limiting the transducer bandwidth first observed in 2002 were successful, resulting in a transducer with a clean frequency response and an effective coupling factor of 0.85. The measured transducer admittance was in nearly exact agreement with theoretical predictions. The NUWC in-water tests demonstrated that the single crystal cylinder achieved an admittance bandwidth (based on the Stansfield criterion) of over 100%, while the tuned power factor was 0.8 or more over 2.5 octaves of frequency. Additionally, the transducer produced 12 dB higher source levels than a similarly sized PZT transducer. [Work sponsored by DARPA.]

3:05

2pEA6. Dipole projectors for conformal sonar system applications. Michael D. Gray (Acoust. and Mech. Systems Div., Georgia Tech Res. Inst., Atlanta, GA 30332-0810), Peter H. Rogers, and Gregg D. Larson (Georgia Inst. of Technol., Atlanta, GA 30332-0405)

Dipole projectors are being investigated for use in active sonar arrays on compliantly coated, low input impedance hulls. When the projector standoff d is acoustically small ($k_0 d < 1$), the direct and hull-reflected signals add in phase in the far field, leading to an increase in signal

strength on the order of 2 relative to a free field dipole. By contrast, the direct and reflected signals for monopole transducer will have nearly opposite phase, and the monopole will have a reduction in total signal strength on the order of $2k_0 d$. Ideally, both transducer types have cosine directivities in-situ, although the monopole transducer response can be more strongly influenced by hull elastic contributions. Designs for two simple realizations of a dipole projector will be presented, along with beam pattern measurements made with prototype transducers at the Georgia Tech water tank facility.

3:20

2pEA7. A novel underwater acoustic transmitter. Chung Chang and Richard Coates (Schlumberger-Doll Res., 36 Old Quarry Rd., Ridgefield, CT 06877)

A strong wide bandwidth low-frequency acoustic transmitter was designed and built to operate in the high-pressure environment of 20 000 psi and temperature over 200 °C inside an oil well. The same design idea can be easily adapted to the naval application in order to use it in the deep ocean environment. This transmitter is pressure balanced and does not use conventional piezoelectric material. Therefore, there is no performance degradation with changing pressure and temperature. This new transmitter uses impact force onto an acoustic resonator to generate sound. The impact force is generated by an electrical hammer. The design is simple, robust and the radiated acoustic wave fields are as repeatable as the piezoelectric source. In addition this new transmitter does not require an expensive amplifier to operate. A video clip will be shown to demonstrate the strength of the source. Further research was done to understand the impact physics and its controlling parameters. Using signal deconvolution technique one can measure the impact force function. A peak impact force of a few thousand pounds by the hammer was measured.

3:35

2pEA8. 3D spatial sampling with a cylindrical multibeam sonar array. Daniel S. Brogan and Christian P. de Moustier (Ctr. for Coastal and Ocean Mapping, Univ. of New Hampshire, 24 Colovos Rd., Durham, NH 03824-3525, daniel.brogan@unh.edu)

Various beam pair combinations can be formed with cylindrical multibeam sonar arrays to obtain a 3D spatial sampling of a patch of seafloor for each ping. This capability is explored with a 286 deg sector cylindrical array transmitting a stepped FM pulse over a 243 deg vertical fan beam centered on nadir and receiving with twenty-seven beam pairs, symmetrically steered about nadir in the fore-aft direction and spaced at 7.16 deg intervals across track. Conventional conjugate product techniques yield two across-track profiles of 26 soundings each (52 phasors total) per ping. However, by combining one or two phasors along-track and one, two or three phasors across-track, soundings can be obtained at the spatial bisectors of the angles between the 52 phasors. This yields three profiles consisting of 51 soundings each, for a total of 153 soundings. This effectively creates a 3D patch of detected seafloor for each ping, which may overlap with the patches from adjacent pings depending on the sonars attitude changes and altitude above the seafloor. [Work supported by NRL-SSC grant N00173-00-1-G912 and NOAA grant NA170G228.]

2pEA9. Techniques in piezoelectric transducer health monitoring. Eli M. Hughes I, Karl M. Reichard, and Tom B. Gabrielson (Penn State Appl. Res. Lab., 3075 Research Dr., CATO Park State College, PA 16801)

For many high intensity underwater applications, the piezoelectric elements within a transducer can be placed under high thermal, electrical and mechanical stress. One particular area of piezoelectric transducer research that has received little attention is *in-situ* health monitoring and failure prediction of the piezoelectric elements. While some studies have been performed in the theoretical mechanics of fractures due to device stress, there is a lack of practical information about piezoelectric health

monitoring and failure prediction. This study was directed at providing practical information about a transducer's electrical characteristics as it is stressed such that failure predictions can be made. The two major failure modes examined in this study were electro-mechanical stress and thermal depolarization. A large part of this research was devoted to the development of novel low-cost electronics that permit measurement of the broad band electrical impedance of the device from 0 to 50 kHz in a small time aperture. The new measurement techniques allow for a large amount of visibility in how a transducer's electrical impedance changes before and during failure. This added visibility provides information that can be used to accurately predict if a transducer is failing and what mechanism is causing the failure.

TUESDAY AFTERNOON, 17 MAY 2005

REGENCY F, 1:00 TO 4:30 P.M.

Session 2pNS

Noise: Environmental Noise and Noise Criteria

John Erdreich, Chair

Ostergaard Acoustical Associates, 200 Executive Dr., West Orange, NJ 07052

Contributed Papers

1:00

2pNS1. Blast noise impacts on sleep. Edward T. Nykaza and Larry L. Pater (Eng. Res. and Development Ctr., Construction Eng. Res. Lab., 2902 Farber Dr., Champaign, IL 61822)

Firing large guns during the hours of darkness is essential to combat readiness for the military. At the same time most people are particularly sensitive to noise when sleeping or trying to fall asleep. Laboratory studies done by Griefahn [J. Sound and Vib. **128**, 109–119 (1989)] and Luz [see Luz *et al.*, ERDC/CERL, TR-04-26 (2004)] suggest that a time period at night may exist where people are more tolerant to large weapon impulse noise (blast noise) and therefore, are less likely to be awakened from noise events. In the fall of 2004, a field study was conducted around a military installation to determine if such a time period(s) exists. Noise monitors were set up inside and outside of residents homes to record noise levels from live military training activities and actimeters were worn by participants sleeping their natural environment to measure sleep disturbance and awakening. The method and results of this study will be presented. [Work supported by US Army Engineer Research and Development Center CERL.]

1:15

2pNS2. Development of metrics to identify military impulse noise. Jeffrey S. Viperman (Dept. of Mech. Eng., Univ. of Pittsburgh, 648 Benedum Hall, Pittsburgh, PA 15261)

Urban encroachment of military bases continues to be a serious problem, affecting training exercises and overall military readiness. The military currently performs noise monitoring around bases to ensure that environmental conditions are favorable for tests or exercises. Despite refinements, current monitoring systems still suffer from occasional false positives and rely heavily on human interpretation. The long range goal of this work is to create a more accurate and autonomous noise classifier. The specific aims of the work are to create a library of recorded military impulse noise sources and to develop, test, and refine the noise classifier algorithms. Representative, high quality field measurements of various noise sources (25 mm and larger) will be conducted at a few military bases under varying environmental conditions. Measurements of wind noise and thunder will also be recorded. The software will be developed in MATLAB and will employ several common impulse noise metrics used for hearing conservation, including rise, peak, and fall times, crest factor, A-

and B-durations, L_{pk} and kurtosis, among others. A neural network will be trained to discern between the recorded military noise sources and naturally occurring noise based on these metrics. Preliminary results from the measurements and analysis will be presented. [Work supported by SERDP.]

1:30

2pNS3. Highway noise levels in a suburban environment under inversion conditions. James Chambers (Dept. of Mech. Eng., Univ. of Mississippi, University, MS 38677), Hugh Saurenman (ATS Consultants Inc., Los Angeles, CA 90017), Robert Bronsdon, Louis Sutherland (Simi Valley, CA 93063), Ken Gilbert, Roger Waxler, and Carrick Talmadge (The Univ. of Mississippi, University, MS 38677)

Noise levels were measured in Scottsdale, AZ during a two week period in March, 2004 to identify the reasons for increased noise levels near a major highway during early morning hours. The noise levels were accompanied by meteorological measurements as well as traffic counts to fully describe the problem. The noise levels were measured in one-third octave bands and ranged 100 ft to 2620 ft (30–800 m) from the highway and included data on both sides of the highway. The meteorological data indicated inversion conditions or downward refraction during the times of interest and model results from a Parabolic Equation (PE) calculation indicated good results with the data. The data and model indicated an approximately 10–15 dB increase in levels during inversion conditions which rapidly transitioned to neutral and lapsed conditions shortly after sunrise. The results of the modeling effort as well as the data will be presented. [Work supported by Arizona Dept. of Transportation.]

1:45

2pNS4. Determination of noise descriptors and criteria for pyrotechnic noise sources. Weixiong Wu (AKRF, 117 East 29th St., New York, NY 10016)

A noise study was conducted to determine appropriate noise descriptors and criteria for assessing pyrotechnic noise sources. The study was carried out to support an environmental impact statement (EIS) that defined sensitive land uses adjacent to reservoirs in New York City area, where potential noise impacts from avian dispersion measures would occur. The pyrotechnic techniques defined as impulsive noise sources are

among the avian dispersion measures that would be used at the reservoirs. Determining appropriate noise descriptors and criteria was critical to the EIS because of the distinctive sound characteristics of pyrotechnic impulse noise sources, the lack of published literature on assessing them, and the absence of corresponding noise regulations. Noise descriptors and criteria used for EIS in the United States were investigated, and noise measurements for pyrotechnic noise sources and some impulsive noise sources were also performed. The study results demonstrate that *C*-weighted DNL is an appropriate descriptor for assessing noise impacts from the pyrotechnics based on the U.S. Army Environmental Noise Management Program criteria, and peak hour *A*-weighted $Leq(1)$ is a suitable noise descriptor for determining noise impacts for avian dispersion measures, including the pyrotechnics, based on the New York City Environmental Quality Review criteria.

2:00

2pNS5. A study of automobile exhaust noise preferences. Jay B. Haire, Melinda J. Carney, and Dominique J. Cheenne (Dept. of Audio Arts & Acoust., Columbia College Chicago, Chicago, IL 60605)

A study was conducted to investigate the relationship between preferences in automobile exhaust noise and the demographic factors of a listening jury. Noise samples of four different vehicles were recorded at idle as well as at 3000 RPM, and 1/3 octave sound spectra were acquired simultaneously. The recordings were presented to the jury using headphones and a preference survey was administered. Zwicker loudness was computed for all samples. Demographic factors such as gender, age, current and future vehicle ownership, were correlated to listening preferences, and unforeseen results were found, especially in regards to sport utility vehicles (SUV).

2:15

2pNS6. Sound-quality analysis of sewing machines. James Chatterley, Andrew Boone, Jonathan Blotter (Dept. of Mech. Eng., Brigham Young Univ., 435 CTB, Provo, UT 84602), and Scott Sommerfeldt (Brigham Young Univ., Provo, UT 84602)

Sound quality analysis procedure and results for six sewing machines ranging from entry level to professional grade will be presented. The procedure consisted of jury-based listening tests and quantification of sound quality using standard metrics. The procedures and analysis of the jury testing will be presented and discussed. The correlation between the quantitative metrics and the qualitative jury results will be presented. Sound localization scans, using near field acoustic holography techniques with accompanying results, performed in order to determine machine sound hot spots and possible sources for undesired sounds, will also be presented. Proposed modifications to machine structure in order to alter machine sound signature into a more sensory pleasant sound will also be presented.

2:30

2pNS7. Criteria for multiple noises in residential buildings using combined rating system. Jin Yong Jeon, Jong Kwan Ryu, and Young Jeong (School of Architectural Eng., Hanyang Univ., Seoul 133-791, Korea)

Multiple residential noises such as floor impact, air-borne, bathroom, drainage, and traffic noises were classified using a combined rating system developed from a social noise survey and auditory experiments. The effect of individual noise perception on the evaluation of the overall noise environment was investigated through a questionnaire survey on annoyance, disturbance, and noise sensitivity. In addition, auditory experiments were undertaken to determine the allowable sound pressure level for each residential noise source and the percent satisfaction for individual noise levels. From the results of the survey and the auditory experiments, a combined rating system was developed and annoyance criteria for multiple residential noises were suggested.

2pNS8. Noise impact on community: A case study for power generation facility. Yong Ma, Jonathan Chui, and Salem Hertil (ATCO Noise Management, 1243 McKnight Blvd. NE, Calgary, AB, Canada)

Power generation plant will make noise impact on the surrounding communities and cause noise complaints from the residences. Noise mitigation treatment for plant is required to achieve the specified noise regulations. In this paper, a case study of the noise control design for a power generation facility is presented. Major noise sources included five engines and generators, five gas conditioning skids, five radiator coolers, and other accessory equipment. The acoustic modeling software Cadna/A was used to predict the noise contributions from sources and assess the noise impact on the nearby communities. During the acoustic modeling, alternative noise mitigation measures underwent two specific investigations before they were chosen as a noise solution recommendation. The first was to determine the technical feasibility of attenuating the source equipment. The second was to perform a cost benefit analysis, necessary to find the most cost-effective solution. For example, several acoustic wall and roof assemblies were entered into the acoustic model and the acoustic performance of the ventilation system was varied until we were able to achieve the most economical acoustic solution.

3:00

2pNS9. Multi-channel active noise control on an axial fan using variable loads. Connor Duke and Scott Sommerfeldt (Brigham Young Univ., N283 ESC Provo, UT 84602)

A multi-channel active noise control system using a mock computer enclosure has been shown to produce significant reduction in tonal noise produced by an axial fan. The mock computer enclosure creates an impedance which influences the fan operation. For this system to be used commercially, it should be versatile enough to be used in a variety of enclosures and environments. Various enclosures and environments can be simulated using a fan plenum to create different back pressures on the system. The effects of different loads seen by the system on the performance of the system will be discussed. The performance of the system can be characterized in either a reverberant or anechoic environment, and the results so obtained will be presented.

3:15

2pNS10. Low-frequency noise and air vibration generated by a simple cycle gas turbine installation. Chris Giesbrecht and Salem Hertil (ATCO Noise Management, 1243 McKnight Blvd NE, Calgary, AB, Canada T2E 5T1, Chris.Giesbrecht@atconoise.com)

Standard noise regulations, measurement techniques, and acoustic treatments can fail to address energy emitted by certain problem gas turbine installations in the form of low-frequency noise and air vibrations. Survey and treatment of the entire acoustic environment is necessary to diagnose and solve these annoyance problems. The presence of ground-borne vibrations, infrasound, and noise in the lower octave bands complicates environmental noise assessment and limits mitigation options. Variability in the perception, effects, and annoyance level of low-frequency noise and air vibrations prevents this relatively common problem from being well understood by industry. Particular attention must be paid to propagation paths when the dynamic range of air-borne and ground-borne vibrations overlaps. This paper is the case study of a 2002 analysis, by ATCO Noise Management, of an installation consisting of 3–120 MW combustion generator turbines causing such low-frequency noise and air vibrations.

2pNS11. Global active control of broadband noise from small axial cooling fans. Matthew J. Green and Scott D. Sommerfeldt (Dept. of Phys. and Astron., Brigham Young Univ., N-281 ESC, Provo, UT 84602, green.mattgreen@gmail.com)

The filtered-x LMS algorithm has previously been used for feed-forward control of the tonal noise of a small axial cooling fan. This system consists of four actuator-sensor pairs surrounding the small fan. The ideal placement of these elements has been previously determined by Gee and Sommerfeldt [Gee and Sommerfeldt, *J. Acoust. Soc. Am.* **115**, 228–236 (2004)]. With success in reducing tonal noise to about the same level as broadband noise it now becomes desirable to control the broadband noise of the small axial cooling fan. This is accomplished with active feedback control. The performance of the broadband control system will be reviewed, and the results of a hybrid (feed-forward and feedback) system for overall reduction of cooling fan noise will be discussed. Ideal methods and configurations for feedback control will also be discussed.

2pNS12. A predictive noise study regarding the proposed Cincinnati Municipal Airport expansion. Grant E. Limberg, Melinda J. Carney, and Dominique J. Cheenne (Dept. of Audio Arts & Acoust., Columbia College Chicago.)

A noise study was performed on the landing path of runway 21L of Cincinnati Municipal Airport to determine the effect of the runway expansion project set to begin in 2005. Sound pressure levels were acquired along the landing path branches to evaluate the eight-hour Leq and the test data were compared to those predicted by the Integrated Noise Model (INM) software. The test data line up to the modeled results within 3.5 dB, an amount that can be explained by weather patterns and other environmental noise variables. The model shows that the 65 dB Day/Night Level (DNL) contour area could negatively affect 46 residents near the airport. Suggestions for the expansion plans include a detailed review of all relevant noise issues using an appropriate computer model.

2pNS13. An outdoor noise propagation study to predict the effect of a power plant expansion. Philip J. Brasovan, Melinda J. Carney, and Dominique J. Cheenne (Dept. of Audio Arts & Acoust., Columbia College Chicago, 600 S. Michigan Ave., Chicago, IL 60605)

The results of an outdoor noise propagation model using CadnaA were compared to test data obtained on-site. The subject property is the central utility plant of a hospital located in Milwaukee, scheduled to be expanded with the addition of cooling towers. The modeled area was 400 m squared with a resolution grid of 2 m squared. The model was used to validate the observed test data as well as to predict the anticipated noise levels at completion of the expansion. A total of 11 points were investigated and the predicted data were found to match the test values within 2 dB at many locations. The data from the model show that the anticipated noise levels at the East property line will exceed those mandated by local ordinances by 3 dB. The model also predicts that the addition of a three meter absorbing barrier and the use of reduced noise fans for the six cell cooling system will bring the overall noise level from the system into compliance.

2pNS14. Is Lighthill's acoustic analogy still relevant? Werner Richarz (Aercoustics Eng. Ltd., 50 Ronson Dr., Ste. 165, Toronto, ON, Canada M9W 1B3)

More than fifty years after Lighthill's seminal paper, the acoustic analogy appears to be out of favor, even though the theory has been able to cope with all aspects of jet noise. This review examines the principal features of Lighthill's acoustic analogy through the lens of the self-and shear noise theory (one of several possible formulations), introduced by Proudman and extended by Ribner and others. One to infer a great deal about the sound field of a jet with a minimum of information about the details of the flow. For example, the well-known U8 scaling law is readily extended to predict the variation of the far-field power spectra of jet noise. The validity has been demonstrated by applying the prediction algorithm to measured jet noise data. Correlations of the source terms and the far-field sound have been measured and shown to exhibit closure. Subtle features such as two-point correlations of far-field sound pressures are accurately described. Recently, Ribner has proposed a means by which the elusive refraction due to flow and temperature gradients can be dealt with in a consistent manner. Predictions based thereon agree quite well with measurement, even for small angle from the jet axis.

Session 2pPA

Physical Acoustics, Engineering Acoustics and Noise: Infrasound: A New Frontier in Monitoring the Earth II

Wayne N. Edwards, Cochair

Dept. of Earth Sciences, Univ. of Western Ontario, 1151 Richmond St., London, ON N6A 5B7, Canada

David E. Norris, Cochair

BBN Technologies, 1300 North 17th St., Arlington, VA 22209

Contributed Papers

1:30

2pPA1. Assessment of atmospheric models for tele-infrasonic propagation. Mihan McKenna (Dept. of Geological Sci., Southern Methodist Univ., P.O. Box 750395, Dallas, TX 75275-0395, smckenna@smu.edu) and Sylvia Hayek (Natural Resources, Canada)

Iron mines in Minnesota are ideally located to assess the accuracy of available atmospheric profiles used in infrasound modeling. These mines are located approximately 400 km away to the southeast (142) of the Lac-Du-Bonnet infrasound station, IS-10. Infrasound data from June 1999 to March 2004 was analyzed to assess the effects of explosion size and atmospheric conditions on observations. IS-10 recorded a suite of events from this time period resulting in well constrained ground truth. This ground truth allows for the comparison of ray trace and PE (Parabolic Equation) modeling to the observed arrivals. The tele-infrasonic distance (greater than 250 km) produces ray paths that turn in the upper atmosphere, the thermosphere, at approximately 120 km to 140 km. Modeling based upon MSIS/HWM (Mass Spectrometer Incoherent Scatter/Horizontal Wind Model) and the NOGAPS (Navy Operational Global Atmospheric Prediction System) and NRL-GS2 (Naval Research Laboratory Ground to Space) augmented profiles are used to interpret the observed arrivals.

1:45

2pPA2. Infrasound radiation of cyclones. Konstantin A. Naugolnykh (NOAA/ETL/Zeltech, 325 Broadway, Boulder, CO 80305)

Tropical cyclones produce strong perturbations of atmosphere and the ocean surface accompanied by acoustical radiation. Infrasonic signals in the 0.1–0.5 frequency band can be observed at distances of thousands of miles from the cyclone. The effect of infrasound radiation is connected apparently to the interaction of the counter-propagating sea-surface waves that produces a sound radiation of the doubled frequency of the surface wave oscillation. This radiation has narrow-angle vertical directivity pattern. The essential refraction of radiated infrasound in the atmosphere perturbed by the cyclone leads to trapping of the infrasound by the horizontal atmospheric wave-guide providing its long distance propagation.

2:00

2pPA3. Characteristics of infrasound signals from earthquakes. Paul Mutschlechner and Rodney Whitaker (Los Alamos Natl. Lab., EES-2 MS J577, LANL, Los Alamos, NM 87545)

Analysis of infrasound signals is presented for a set of 31 earthquakes located mostly in the western United States and recorded at infrasound arrays operated by the Los Alamos National Laboratory. By normalizing measured amplitudes for the effects of propagation and distance, a pressure-amplitude versus earthquake magnitude relation is derived. Further analysis showed that the observed variance in this relation is likely due to the variation in source ground motion and variation in upper atmospheric winds. Signal durations can be tens of minutes and azimuth devia-

tions have a mean of three degrees (measured from array data compared to great circle azimuths). Other characteristics of the infrasound data will be discussed. Our analysis of the observed data combined with surface ground motion accelerations, led to estimates of the minimum surface accelerations needed to generated infrasound signals measured at distant arrays.

2:15

2pPA4. An analysis of seismic and acoustic signals from the June 3, 2004 Washington state bolide. Stephen Arrowsmith, Michael Hedlin (IGPP, Scripps, Univ. of California, San Diego, La Jolla, CA 92093-0225), Lars Ceranna (BGR, Hannover 30625, Germany), and Wayne Edwards (Univ. of Western Ontario, Canada, N6A 3K7)

On June 3rd, 2004 a spectacular bolide was reported over British Columbia, Washington, Oregon and Idaho. In addition to eyewitness accounts and video recordings, the event was recorded on a number of seismometers in the Pacific Northwest Seismograph network and at infrasound arrays in Washington State and California. Using the NRL-G2S atmospheric model for the time of the event, source locations have been determined with seismic and acoustic signals separately. Estimates of the yield and acoustic efficiency of the explosion have also been determined. By simulating the propagation of infrasound through the atmosphere, the arrival of discrete acoustic phases at the infrasound arrays has been modeled.

2:30

2pPA5. Acoustic observations of large earthquakes and associated phenomena. Catherine de Groot-Hedlin (Scripps Inst. of Oceanogr., UCSD, 9500 Gilman Dr., La Jolla, CA 92093-0225, chedlin@ucsd.edu) and David McCormack (Geological Survey of Canada, Ottawa, ON, Canada K1A 0Y3)

Recent global events have renewed interest in acoustic observations of large earthquakes and associated phenomena such as landslides and tsunamis. In addition, the proliferation of global networks of atmospheric acoustic (infrasound) and hydroacoustic observing systems mean that signals from such events are captured in close to real-time more frequently and with better resolution than ever before. Such data provide a useful complement to more traditional seismological observations of such events. In this paper, recent acoustic and hydroacoustic observations from large earthquake events are presented. As well as the scientific insights provided by use of such monitoring systems, operational applications useful for hazard mitigation or hazard alerting are considered. Current plans to develop observing systems to perform systematic studies of areas susceptible to earthquake and tsunami hazard offshore western Canada, and the use of such proposed systems to make systematic studies of hydroacoustic signals from seismic events, underwater slumping and landslides are described.

2:45

2pPA6. High trace-velocity events associated with auroral activity in Alaska. Charles R. Wilson and John V. Olson (Geophysical Inst., Univ. of Alaska, Fairbanks, AK 99775)

Data from the CTBT/IMS station 153US at Fairbanks, AK covering 2003 and 2004 have been surveyed for high-correlation, high trace-velocity events. Although such events have been detected at all times of the day the distribution shows a strong maximum in the morning hours (near 15UT) and a weaker maximum in the evening hours (near 5UT). Twenty intervals containing high trace-velocity events from 2003 and 37 intervals from 2004 have been studied in detail. The intervals containing high trace-velocity events can last from a few hours up to 10 hours. All-sky video records show that the high trace-velocity events are associated with stable, pulsating auroral patches that occur during the recovery phase of substorms. This talk will review the characteristics of the high trace-velocity events and the associated auroral video where available.

3:00–3:30 Break

3:30

2pPA7. Source localization of non-stationary acoustic data using time-frequency analysis. Jack Stoughton (NASA-Langley Res. Ctr., MS 488, Hampton, VA) and William Edmonson (Hampton Univ., Hampton, VA 23668)

An improvement in temporal locality of the generalized cross-correlation (GCC) for angle of arrival (AOA) estimation can be achieved by employing 2-D cross-correlation of infrasonic sensor data transformed to its time-frequency (TF) representation. Intermediate to the AOA evaluation is the time delay between pairs of sensors. The signal class of interest includes far field sources which are partially coherent across the array, nonstationary, and wideband. In addition, signals can occur as multiple short bursts, for which TF representations may be more appropriate for time delay estimation. The GCC tends to smooth out such temporal energy bursts. Simulation and experimental results will demonstrate the improvement in using a TF-based GCC, using the Cohen class, over the classic GCC method. Comparative demonstration of the methods will be performed on data captured on an infrasonic sensor array located at NASA Langley Research Center (LaRC). The infrasonic data sources include Delta IV and Space Shuttle launches from Kennedy Space Center which belong to the stated signal class. Of interest is to apply this method to the AOA estimation of atmospheric turbulence. [Work supported by NASA LaRC Creativity and Innovation project: Infrasonic Detection of Clear Air Turbulence and Severe Storms.]

3:45

2pPA8. Multiparameter studies of surf infrasound. Milton Garces, David Fee, Pierre Caron, Claus Hetzer (Infrasound Lab., Univ. of Hawaii, Manoa, 73-4460 Queen Kaahumanu Hwy., #119, Kailua-Kona, HI 96740-2638), Jerome Aucan, Mark Merrifield (Univ. of Hawaii, Manoa, Honolulu, HI 96822), Robert Gibson, and Joydeep Bhattacharyya (BBN Technologies, Arlington, VA 22209)

Infrasound stations on islands or near coastlines routinely detect signals associated with breaking ocean waves. Although the source mechanisms of these infrasonic surf signals are not well understood, they might provide useful insight into coastal processes and permit an assessment of

wave energy distribution in the littoral zone. Near-shore infrasonic arrays, ocean bottom sensors, an infrared imager, and a video camera were deployed on rocky and sandy coastlines in Hawaii during the 2004–2005 Winter swell season, with the aim of establishing a relationship between the directional ocean swell height and the infrasonic source distribution, intensity, and spectral content. The cameras targeted acoustically active regions for selected time periods to associate the timing of the infrasonic signal arrivals with the breaking and dissipation of ocean wave sets. A wide range of sea states and weather conditions were captured during two separate deployments. The results of our multiparameter analyses and preliminary comparisons of our observations with mesoscale ocean wave model predictions will be presented. [Work supported by the Defense Advanced Research Projects Agency. Support of the State of Hawaii is acknowledged.]

4:00

2pPA9. Multiple array infrasound observations in the Netherlands. Láslo Evers and Hein Haak (Royal Netherlands Meteorological Inst. (KNMI), P.O. Box 201, 3730 AE De Bilt, The Netherlands, evers@knmi.nl)

Infrasound in the Netherlands is observed with multiple arrays. The current architecture consists of 4 arrays. These arrays vary in aperture from 30 to 1500 m while the number of microbarometers ranges from 6 to 16 per array. Continuous observation of infrasound implies a huge number of automatic detections per year. In other words, thousands of coherent waves cross the arrays each year. Although a lot of sources can be identified, there are still unknown sources having a coherent infrasonic signature. Among the identified sources are: supersonic airplanes, bolides, severe weather, oceanic waves, volcano explosions, and military activity. Current efforts in source identification concentrate on the integration of infrasonic and seismic data. Furthermore, atmospheric characteristics are included such as the state of the boundary layer and propagation conditions in the higher atmospheric, i.e., stratospheric winds and temperatures. In this presentation the influence of the atmosphere on the detection capability will be shown. Furthermore, the detailed analysis of specific sources shows the enormous potential of infrasound as atmospheric probe and monitoring technique.

4:15

2pPA10. Status report on the establishment of the CTBTO IMS infrasound network. Thomas L. Hoffmann (CTBTO IMS, Vienna Intl. Ctr., P.O. Box 1200, A-1400 Vienna, Austria)

Steady progress has been made in the establishment of the CTBTO IMS infrasound monitoring network. To date 86% of the site surveys for 60 infrasound stations in the network have been completed, 50% of the stations are transmitting continuous data to Vienna, and 40% of the stations have been certified. While the global distribution pattern of infrasound stations transmitting data to Vienna is still disperse, regional networks begin to form in North and South America as well as in the Australian and South African regions. This presentation will focus on an overview of recent progress made in the establishment of the global infrasound network, and also present some of the challenges and difficulties encountered in this program.

Session 2pPP

Psychological and Physiological Acoustics: Psychoacoustics: In Memory of Søren Buus

Bertram Scharf, Cochair

Northeastern Univ., Psychology, Boston, MA 02115-5096

Rhona P. Hellman, Cochair

*Northeastern Univ., Speech Language Pathology and Audiology, 360 Huntington Ave., Boston, MA 02115***Chair's Introduction—1:00*****Invited Papers***

1:05

2pPP1. Søren Buus. Thirty years of psychoacoustic inspiration. Torben Poulsen (Oersted-DTU, Acoust. Technol., Tech. Univ. of Denmark, DK 2800 Lyngby, Denmark, tp@oersted.dtu.dk)

Søren Buus did his MSc at the Acoustics Laboratory, Technical University of Denmark (DTU), in 1975 on the topic headphone calibration. He showed the importance of reliable reference values for psychoacoustic research and Søren was a great inspiration for my work [Scand. Audiol. **20**, 205–207 (1991); **27**, 105–112 (1998)]. Already from the seventies, temporal integration of loudness has been a major topic in the collaboration with Søren [Buus *et al.*, *J. Acoust. Soc. Am.* **105**, 3464–3480 (1999)] and the measurements of temporal integration over a wide range of presentation levels led to the important finding about the shape of the loudness function [Buus *et al.*, *J. Acoust. Soc. Am.* **100**, 669–680 (1997)]. Søren talked about the importance of the psychophysical procedure and the influence from the procedure on the results [Buus, *Proceedings 19th Danavox Symposium* (2001), pp. 183–226]. The goal was to obtain reliable, unbiased, and precise results. An overview of some of the above investigations will be presented together with recent results from a MSc project on headphone calibration of short duration sounds for ABR measurements.

1:30

2pPP2. Søren Buus' contribution to speech intelligibility prediction. Hannes Müsch (Sound ID, Palo Alto, CA 94303, hmuesch@soundid.com) and Mary Florentine (Northeastern Univ., Boston, MA 02115)

In addition to his work in psychoacoustics, Søren Buus also contributed to the field of speech intelligibility prediction by developing a model that predicts the results of speech recognition tests [H. Müsch and S. Buus, *J. Acoust. Soc. Am.* **109**, 2896–2909 (2001)]. The model was successful in test conditions that are outside the scope of the Articulation Index. It builds on Green and Birdsall's concept of describing a speech recognition task as selecting one of several response alternatives [in D. Green and J. Swets, *Signal Detection Theory* (1966), pp. 609–619], and on Durlach *et al.*'s model for discriminating broadband sounds [*J. Acoust. Soc. Am.* **80**, 63–72 (1986)]. Experimental evidence suggests that listeners can extract redundant, independent, or synergistic information from spectrally distinct speech bands. One of the main accomplishments of the model is to reflect this ability. The model also provides for a measure of linguistic entropy to enter the intelligibility prediction. Recent model development has focused on investigating whether this measure, the cognitive noise, can account for the effects of semantic and syntactic context. This presentation will review the model and present new model predictions. [Work supported by NIH grant R01DC00187.]

1:55

2pPP3. Borrowing a trick from Søren Buus: Application of a simple quantitative model to data obtained from an impaired population. Robert P. Carlyon (MRC Cognition and Brain Sci. Unit, 15 Chaucer Rd., Cambridge, CB2 2EF, England)

One of the many things that Søren taught me was to apply a simple quantitative model to psychophysical data, even when studying impaired populations, renowned for inter-listener variability. Here I present a recent application of this approach, applied to cochlear implant patients' detection thresholds for electrical stimulation. Most implants stimulate the nerve with trains of biphasic pulses, in which a rectangular pulse of one polarity is followed, with a short inter-phase gap (IPG), by an equal pulse of opposite polarity. Together with van Wieringen and colleagues in Leuven, Belgium, we have shown that psychophysical thresholds in humans drop as IPG increases to at least 2900 microseconds implying a much longer time constant than predicted from physiological studies on animals. By applying a simple model, we showed that this long time constant is quantitatively consistent with the threshold drop observed for electrical sinusoids as frequency is lowered to about 100 Hz. The model involves passing the electrical waveform through a low-pass filter and calculating the RMS output. Despite its simplicity, the model successfully predicts the results of several novel manipulations, and helps constrain explanations of the physiological basis of the observed time constants.

2pPP4. Sequential grouping constraints on across-channel auditory processing. Andrew J. Oxenham (Res. Lab. of Electron., MIT, Cambridge, MA 02139, oxenham@mit.edu) and Torsten Dau (Tech. Univ. of Denmark, DK-2800 Lyngby, Denmark)

Søren Buus was one of the pioneers in the study of across-channel auditory processing. His influential 1985 paper showed that introducing slow fluctuations to a low-frequency masker could reduce the detection thresholds of a high-frequency signal by as much as 25 dB [S. Buus, *J. Acoust. Soc. Am.* **78**, 1958–1965 (1985)]. Søren explained this surprising result in terms of the spread of masker excitation and across-channel processing of envelope fluctuations. A later study [S. Buus and C. Pan, *J. Acoust. Soc. Am.* **96**, 1445–1457 (1994)] pioneered the use of the same stimuli in tasks where across-channel processing could either help or hinder performance. In the present set of studies we also use paradigms in which across-channel processing can lead to either improvement or deterioration in performance. We show that sequential grouping constraints can affect both types of paradigm. In particular, the perceptual segregation of off-frequency from on-frequency components, using sound sequences preceding or following the target, leads to results similar to those found in the absence of the off-frequency components. This suggests a high-level locus for some across-channel effects, and may help provide a functional distinction between within- and across-channel mechanisms. [Work supported by NIH R01DC03909 and Danish Research Council.]

2:45–3:00 Break

3:00

2pPP5. Tone-burst otoacoustic emissions and loudness. Michael Epstein (Inst. of Hearing, Speech and Lang., Comm. Res. Lab, and Comm. and DSP Ctr., ECE Dept. (440 DA), Northeastern Univ., Boston, MA 02115, mepstein@ece.neu.edu) and Mary Florentine (Northeastern Univ., Boston, MA 02115)

Several models of cochlear mechanics lead to the conclusion that the amplitude of tone-burst otoacoustic emissions (TBOAEs) is proportional to basilar-membrane (BM) motion. Buus and Florentine [Fechner Day 2001 (Pabst, Berlin), 236 (2001)] showed that the square of loudness, derived from measurements of spectral and temporal integration, closely matched basilar-membrane velocity at the best frequency. Buus *et al.* [Physiological and Psychophysical Bases of Auditory Function, 373 (2001)] examined the relationship between distortion-product otoacoustic emissions (DPOAEs) and BM motion. Follow-up work by Epstein *et al.* [*J. Acoust. Soc. Am.* **117**, 263 (2005)] showed in a direct comparison that TBOAEs and psychoacoustical measures, of loudness and of pulsation threshold, on six subjects with normal hearing led to very similar estimations of basilar-membrane motion. These outcomes suggest that otoacoustic emissions could serve as an excellent tool—one that is objective, non-invasive, and rapid—for estimating in subjects with normal hearing both relative basilar-membrane motion and relative loudness. [Work supported by NIH/NIDCD Grant R01DC02241.]

3:25

2pPP6. Induced loudness reduction: A review. Bertram Scharf (Psychol. Dept., Northeastern Univ., Boston, MA 02115), Eva Wagner, and Bärbel Nieder (Northeastern Univ., Boston, MA 02115, scharf@neu.edu)

Under appropriate stimulus conditions, a tone may decline in loudness the equivalent of 10 dB and more when preceded by a stronger tone. This induced loudness reduction or ILR was uncovered, indirectly, in a large number of studies by L. E. Marks and his associates [e.g. Marks, *J. Exp Psychol HPP* **20**, 382–396 (1994)]. Those studies seemed to suggest that ILR required that tones be presented at two widely separated frequencies over a relatively large range of SPLs. Although part of the measured loudness changes seemed to stem from response biases, recent studies show unequivocally that the reduction in loudness is mostly sensory and that tones need be presented at only a single frequency and at two levels some 10 to 20 dB apart. The present paper puts together what is known about the dependence of ILR on signal frequency, level, duration, temporal relations, and hearing loss. The role of ILR in various other psychoacoustical phenomena such as induced loudness adaptation and loudness enhancement is reviewed. Much of this knowledge was obtained in direct and indirect collaboration with Søren Buus. [Work supported by NIH/NIDCD Grant No. R01 DC 02241.]

Contributed Paper

3:50

2pPP7. Extension of Søren Buus's modeling to loudness growth at high frequencies. Rhona P. Hellman (Dept. of Speech-Lang. Path. and Audiol. and Inst. for Hearing, Speech, & Lang. (106A FR), Northeastern Univ., Boston, MA 02115, hellman@neu.edu)

Loudness matching between tones at different frequencies leads to results that imply that loudness functions, which relate loudness to sound pressure level (SPL) are parallel at frequencies from 1 to 10 kHz but not at higher frequencies [Hellman *et al.*, *J. Acoust. Soc. Am.* **109**, 2349 (2001)]. Within the 12.5 to 16 kHz frequency range, the loudness-matching func-

tions are curvilinear in shape being steeper below 60 phons than at higher loudness levels. The higher the frequency, the greater is the decrease in slope above 60 phons. The slope reduction above 60 phons for frequencies from 12.5 to 16 kHz is ascribed to the tone's restricted excitation pattern. A computer model by Søren Buus described in part in Florentine *et al.* [*Modeling Sensorineural Hearing Loss* (Erlbaum, 1997), pp. 187–198] provides a good account of the decrease in the mid-to-high level slope of loudness functions derived at high frequencies from the matching data. The agreement between the model predictions and the empirical results further supports the notion that the rate of loudness growth at moderate-to-high SPLs depends on the upward spread of excitation.

Session 2pSA

Structural Acoustics and Vibration: General Vibration; Excitation, Radiation, and Dampening

Kenneth D. Frampton, Chair

*Dept. of Mechanical Engineering, Vanderbilt Univ., Nashville, TN 37235-1592***Contributed Papers**

2:30

2pSA1. Non-contact mode excitation of small structures in air using ultrasound radiation force. Thomas M. Huber, John C. Purdham (Dept. of Phys., Gustavus Adolphus College, 800 College Ave., St. Peter, MN 56082, huber@gustavus.edu), Mostafa Fatemi, Randall R. Kinnick, and James F. Greenleaf (Mayo Clinic College of Medicine, Rochester, MN 55905)

With the advent of MEMS, modal analysis of small structures is increasingly important. However, conventional excitation techniques normally require contact, which may not be feasible for small objects. We present a non-contact method that uses interference of ultrasound frequencies in air to produce low-frequency excitation of structures. Objects studied included hard-drive HGA suspensions and MEMS devices. The vibration induced by the ultrasound radiation force was varied in a wide range from 0 Hz to 50 kHz. Object motion was detected using a laser vibrometer; measured frequencies agreed with expected values. Also demonstrated was the unique capability to selectively enhance or suppress modes independently. For example, the ratio of the vibrational amplitudes of the 175 Hz first-bending and 1.33 kHz torsional modes of a small cantilever could be changed from in excess of 10:1 to less than 1:10 by shifting the ultrasound modulation phase 90 degrees. Similar changes were obtained for a 3 mm square MEMS mirror in the ratios of vibration amplitude around its two separate axes. Torsional modes of a hard-drive suspension could be selectively enhanced by over a factor of two by moving the ultrasound focus point from near the center to near the edge of the suspension.

2:45

2pSA2. Direction selective structural-acoustic coupled radiator. Hee-Seon Seo and Yang-Hann Kim (NOVIC, KAIST, 373-1 Sci. Town Daejeon-si, Korea)

This paper presents a method of designing a structural-acoustic coupled radiator that can emit sound in the desired direction. The structural-acoustic coupled system is consisted of acoustic spaces and wall. The wall composes two plates and an opening, and the wall separates one space that is highly reverberant and the other that is unbounded without any reflection. An equation is developed that predicts energy distribution and energy flow in the two spaces separated by the wall, and its computational examples are presented including near field acoustic characteristics. To design the directional coupled radiator, Pareto optimization method is adapted. An objective is selected to maximize radiation power on a main axis and minimize a side lobe level and a subjective is selected direction of the main axis and dimensions of the walls geometry. Pressure and intensity distribution of the designed radiator is also presented.

3:00

2pSA3. Development of a directivity controlled piezoelectric transducer for sound reproduction. Magella Bédard and Alain Berry (G.A.U.S., Dept. of Mech. Eng., Université de Sherbrooke, 2500, boul. de l'Université, Sherbrooke, Québec, Canada, J1K 2R1, magella.bedard@usherbrooke.ca)

One of the inherent limitations of loudspeaker systems in audio reproduction is their inability to reproduce the possibly complex acoustic directivity patterns of real sound sources. For music reproduction for example,

it may be desirable to separate diffuse field and direct sound components and project them with different directivity patterns. Because of their properties, poly (vinylidene fluoride) (PVDF) films offer lot of advantages for the development of electroacoustic transducers. A system of piezoelectric transducers made with PVDF that show a controllable directivity was developed. A cylindrical omnidirectional piezoelectric transducer is used to produce an ambient field, and a piezoelectric transducers system, consisting of a series of curved sources placed around a cylinder frame, is used to produce a sound field with a given directivity. To develop the system, a numerical model was generated with ANSYS Multiphysics TM8.1 and used to calculate the mechanical response of the piezoelectric transducer. The acoustic radiation of the driver was then computed using the Kirchoff-Helmoltz theorem. Numerical and experimental results of the mechanical and acoustical response of the system will be shown.

3:15

2pSA4. Progress towards an electro-acoustic resonance technique for determining quantitative material and geometrical properties in high contrast multi-layer elastic structures.. Karl Fisher (7000 E. Ave, Livermore, CA 94551)

High contrast multilayered elastic structures continue to be problematic for ultrasonic inspection. Large acoustic material impedance mismatches, refraction, reverberation, multiple echoes, and high elastic attenuations are just some of the issues one is faced with standard high frequency (1–20 MHz) pulse echo detection and imaging methods. In this presentation, we will present progress towards developing a low frequency resonance technique that operates in the 20 to 70 kHz regime. The technique is based on a direct correlation between the electrical impedance of a standard electro-acoustic transducer and the mechanical loading it experiences when placed in contact with a layered elastic structure. Preliminary experimental and theoretical results are in good agreement.

3:30

2pSA5. A simple model for coupled acoustic-structure resonance in Stratospheric Observatory for Infrared Astronomy. Jerry H. Ginsberg (G. W. Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405)

The Stratospheric Observatory For Infrared Astronomy (SOFIA) is a joint project of NASA and the Deutsches Zentrum für Luft- und Raumfahrt that has mounted a 2.5 m, 20 000 kg infrared telescope on a bulkhead of a specially modified Boeing 747-SP. A large sliding door will expose the observation bay to the exterior flow field at Mach 0.85 and 13 km altitude. In the open configuration the interaction of turbulence vortices generated at the leading and trailing edges of the opening has the possibility of inducing a strong acoustic signal. A concern has been raised that the peak frequencies of such a signal might coincide with the cavity resonances. The present work examines the transfer function for a known source in order to identify the cavity resonances. Simplistic reasoning argues that the worst case would occur if the cavity resonant frequencies are close to structural resonances. However, the structure's impedance is very low at its resonances, which means that the cavity resonant frequencies are shifted from their nominal values. The present work uses a simple one-dimensional waveguide model, in which one end is terminated by a damped single-degree-of-freedom oscillator, to explain the coupled-fluid

structure resonance. The characteristic equation and formulas for the pressure and displacement transfer functions are derived. Analysis of these results leads to some surprising insights regarding the role of a structure's stiffness and mass. [Work supported by the NASA.]

3:45–4:00 Break

4:00

2pSA6. Non-intrusive ground-contacting vibrometer for acoustic/seismic landmine detection. James S. Martin, Gregg D. Larson (Georgia Inst. of Tech., Atlanta, GA 30332-0405), and Waymond R. Scott, Jr. (Georgia Inst. of Tech., Atlanta, GA 30332-0250)

The detection of buried landmines using seismic waves and full wave-field measurements has been demonstrated. The technique requires a sensor that is non-intrusive with sufficient fidelity, reproducibility, and noise immunity for imaging processes. This has been accomplished in the past with non-contact techniques. For reasons of cost and scalability in large arrays, ground-contacting sensors are currently of interest as an alternative to these. A ground-contacting sensor configuration was studied in which an accelerometer was coupled to the ground through a viscoelastic layer with a bias force provided by a soft coil spring. This sensor was found to meet the noise, fidelity, and reproducibility requirements of a seismic landmine detection system operating in a laboratory experimental model. The fidelity of this sensor was found to vary with the bias force because of the nonlinear stiffness of the soil surrogate in the model. This dependence was sufficiently weak that no feedback of the bias force was necessary to reproducibly couple the sensor over a flat surface. The sensor offers the potential benefit of information that was not available from non-contact measurements regarding the in-plane motion of the soil surface. This data may provide additional cues for the detection of buried mines. [Work supported by ONR.]

4:15

2pSA7. Vibro-acoustic response of convected fluid loaded plates. Kenneth Frampton (Dept. of Mech. Eng., Vanderbilt Univ., Nashville, TN 37235)

This work demonstrates the effects that fluid convection has on the vibro-acoustic response of rectangular plates. The effects of fluid flow on the vibration and stability of plates is reasonably well understood. These effects include dramatic structural modal coupling along with static and dynamic instabilities. However, the resulting effect on sound radiation is not so well understood. The presentation will include a description of the fundamental physics associated with a simply supported, vibrating, rectangular plate in an infinite baffle and radiating into a semi-infinite, convected fluid. Then, simulation results will demonstrate the effects of flow-induced modal coupling and acoustic radiation. It will be demonstrated that convection can significantly increase the radiated sound power and dramatically affect the vibro-acoustic response.

4:30

2pSA8. Behavior of the intermediate and exterior layers of a close fitting enclosure surrounding a cylindrical fluid-loaded acoustic source. Joseph Cuschieri (Lockheed Martin MS2, Perry Technologies, 100 East 17th street, Riviera Beach, FL 33404)

Previously, results for the Insertion Loss (IL) of a close fitting enclosure surrounding a submerged (in water) cylindrical acoustic source [Cuschieri, J. Acoust. Soc. Am. **115**, 2537 (2004); **116**, 2521 (2004)] were presented for different intermediate layer materials, different characteristics of the exterior cladding layer material and different source sizes. The IL results were presented as a function of frequency and layer/cladding characteristics. In this presentation, details on the behavior of the intermediate and outside layers enclosing the cylindrical source are presented which demonstrate the coupling between the cylindrical source the exter-

nal acoustic (water) medium. These results help to explain the IL results previously obtained and presented. Furthermore, the scattering characteristics of the cylindrical source with and without the close fitting enclosure are considered. [Work supported by ONR.]

4:45

2pSA9. A focused multi-layered spherical shell with guided wave enhancements. John D. Smith (Dstl. Porton Down, Salisbury, SP4 0JQ, UK) and Duncan P. Williams (Dstl. Winfrith, Dorchester DT2 8WX, UK)

There is a continued need for underwater reflective targets to be used as relocation and navigational aids. Traditionally, fluid-filled thin spherical shells have been used as passive reflectors, similar to cat's eyes, and would usually be filled with liquid chlorofluorocarbons (CFC's). Other options are needed now that the production of CFC's has been restricted. This paper looks at the relationship between a system of multi-layered concentric elastic shells and their target strength in water. The problem is formulated using the global matrix method and is applicable to any system with an arbitrary number of layers. Results show that, by using certain combinations of low loss elastic layers together, the target strength is higher than it would be for any one of the materials used in isolation. This result is explained by geometric focussing on the back surface of the sphere and guided elastic waves that circumnavigate the outer layer or shell and enhance the focussing. The results are compared with calculations and existing experiments for a stainless steel shell filled with CFC.

5:00

2pSA10. Sound radiation of a plate excited by an impact. Physical and perceptual comparisons between numerical and experimental results. Dominique Habault, Florence Demirdjian, Sabine Meunier, and Georges Canevet (CNRS-LMA, 13402 Marseille Cedex 20, France)

The study is concerned with the response of a thin elastic baffled plate, immersed in a fluid, and excited by an impact force. An experiment was run, in which a plate was struck by an impact hammer which provides excitations of very short duration. Various shapes of excitation functions were obtained by using different kinds of hammer heads (rubber, plastic, metal). The resulting acceleration on the plate and sound pressure radiated in the fluid were measured, with an emphasis on the very first portion of the signals (initial 30 ms). The recorded signals were then compared, in the time domain, with theoretical predictions based on expansions in resonance modes of the fluid-loaded plate. These predictions provide a quite accurate description of the experimental data. Auditory tests were also run, to compare the signals obtained from the calculations and the recorded sounds. Subjects were asked to evaluate the dissimilarity between test sounds, using an analog scale graduated from 0 (very similar) to 6 (very dissimilar). The tests allowed to identify the perceptual criteria used by the subjects. The aim of the study is to use these criteria to improve the numerical method of prediction of the sound radiation.

5:15

2pSA11. A variational approach to modeling the vibration of timber joist floors. Colin Fox and Hyuck Chung (Mathematics Dept., Univ. of Auckland, PB 92019, Auckland, New Zealand, fox@math.auckland.ac.nz)

We present a comprehensive variational model for the vibration of timber joist floors and a simple computer algorithm for finite, particularly rectangular, floors. We allow for floor constructions that are typical in the New Zealand context, consisting of edge-supported timber joists with flooring material above, often a suspended ceiling below with absorptive material in the cavity, and a range of joint types such as gluing or nailing. The model and algorithm are structured in such a way that a component can easily be either added to, or removed from, the structure. Hence the configuration may be made progressively more complex from the simplest floor-joists type to floor-joists-ceiling with cavity air and damper-spring connectors.

Session 2pSC

Speech Communication: Cross-Linguistic and Dialectal Studies (Poster Session)

Richard A. Wright, Chair

Dept. of Linguistics, Univ. of Washington, Seattle, WA 98195-4340

Contributed Papers

All posters will be on display from 1:00 p.m. to 4:30 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:00 p.m. to 2:15 p.m. and contributors of even-numbered papers will be at their posters from 2:15 p.m. to 4:30 p.m.

2pSC1. Free choice task effects in cross-linguistic stop perception.

Alexei Kochetov (Dept. of Linguist., Simon Fraser Univ., 8888 Univ. Dr., Burnaby, BC, Canada V5A 1S6, alexei_kochetov@sfu.ca)

This paper examines the identification of stop place and secondary articulation using a free choice task. Russian syllable-initial and syllable-final stops /p pj t tj/ in nonsense utterances were presented to Russian and Japanese listeners ($N=30$). Correct identification rates for place and secondary articulation of the target consonants were determined based on written responses (in Cyrillic or Katakana). Both groups of listeners showed better identification of syllable-initial stops compared to syllable-final stops. Among the consonants, /p/ was identified better, and /pj/ was identified worse than the other stops. Native listeners performed better than non-native listeners. The overall correct identification rates were lower than (yet strongly correlated with) the rates previously obtained with the same stimuli using a forced choice phoneme identification task. The lower identification rates in the current study can be explained in part by the errors involving the segmentation and syllabification of palatalized stops. Thus, the palatal articulation of the syllable-final palatalized /pj/ was often interpreted as independent of the stop (e.g., /tapj api/ rendered as /taj papi/ or /tjap api/). It is concluded that the free choice task can successfully complement the forced choice task, providing additional information about the perception of secondary palatalization. [Supported by SSHRC.]

2pSC2. The preliminary study about neutral tone: Dialect effect between North Official Mandarin speakers in China and Taiwan Mandarin speakers.

Jennifer Li (Dept. of English Lit., Natl. Chiao Tung Univ., 1001 Ta Hsueh Rd., Hsinchu, Taiwan 300, Republic of China, wahaha_jennifer@yahoo.com.tw)

According to general theories, neutral tone is not regarded as an independent tone in Mandarin. Previous research shows that the most important characteristic of the neutral tone is that it does not have a certain target and pitch contour. (Lin and Yang, 1980) Namely, its pitch contour is uncertain, and it is weak in perception level. However, those studies ignore the variant between different dialects. Our study examined the features of the neutral tone between two dialects of Mandarin speakers and aimed at figuring out the dialect difference effect on the pronunciation of the neutral tone. Our subjects were chosen from two groups of Mandarin speakers. One group is from the North Mainland China, and the other is from Taiwan. The experiment was designed with a speak-it-out process. All subjects read a randomized script written in Mandarin, and the whole process was recorded spontaneously. The preliminary result shows that the dialect difference effect actually matters. It shows a tendency that the neutral tone has a certain target in Taiwan Mandarin speakers.

2pSC3. The phonetic rhythm/syntax headedness connection: Evidence from Tagalog.

Sonya Bird, Laurel Fais, and Janet Werker (Univ. of Victoria and Univ. of British Columbia, Canada, sbird@uvic.ca)

Ramus, Nespor, and Mehler [Cognition (1999)] show that the rhythm of a language (broadly: stress- versus syllable- versus mora-timing) results from the proportion of vocalic material in an utterance (%V) and the standard deviation of consonantal intervals (delta-C). Based on 14 languages, Shukla, Nespor, and Mehler [submitted] further argue that rhythm is correlated with syntactic headedness: low %V is correlated with head-first languages (e.g., English); high %V is correlated with head-final languages (e.g., Japanese). Together, these proposals have important implications for language acquisition: infants can discriminate across rhythm classes [Nazzi, Bertoncini, and Mehler, J. Exp. Psych: Human Perception and Performance (1998)]. If rhythm, as defined by %V and delta-C, can predict headedness, then infants can potentially use rhythm information to bootstrap into their languages syntactic structure. This paper reports on a study analyzing rhythm in a language not yet considered: Tagalog. Results support the Shukla *et al.* proposal in an interesting way: based on its %V and delta-C, Tagalog falls between head-first and head-last languages, slightly closer to the head-first group. This placement correlates well with the fact that, although Tagalog is said to be primarily head-first syntactically, head-last phrases are permitted and common in the language.

2pSC4. A comparison of the acoustic characteristics of American English and Cantonese vowels.

Eric Zee (Dept. of CTL, City Univ. of Hong Kong, 83 Tat Chee Ave., Hong Kong, ctlee@cityu.edu.hk)

The study compares the formant frequencies of the American English vowels [i, ε, a, o, u, ɪ, ʊ] (Peterson and Barney, 1952) and [i, ε, a, o, u, ɪ, ʊ] in Cantonese (Zee, 2004). Results of the comparison show that the differences in formant values for the vowels between the two languages vary according to vowel type and gender. Between male speakers of the two languages, the differences in F -values are minimal for [i] and [ɔ]. For [ε, u], the difference between the two groups is mainly in $F2$, with [ε] having a larger $F2$ and [u] a smaller $F2$ for Cantonese speakers. [a] in American English has smaller $F1$ and $F2$ than [a] in Cantonese. [ɪ, ʊ] occupy the position in between the level of [i, u] and level of [ε, ɔ] in the $F1/F2$ plane for American English speakers, and there is a noticeable difference in $F2$ between [ɪ] and [ε] and between [ʊ] and [ɔ]. For Cantonese speakers, [ɪ, ʊ] are on the same level of [ε, ɔ], and the difference in $F2$ is minimal between [ɪ] and [ε] and between [ʊ] and [ɔ]. The differences between female speakers of the two languages will also be presented.

2pSC5. Acoustic characteristics of Korean stops in Korean child-directed speech. Minjung Kim and Carol Stoel-Gammon (Dept. of Speech & Hearing Sci., Univ. of Washington, 1417 NE 42nd St., Seattle, WA 98105-6246)

A variety of cross-linguistic studies have documented that the acoustic properties of speech addressed to young children include exaggeration of pitch contours and acoustically salient features of phonetic units. It has been suggested that phonetic modifications of child-directed speech facilitate young children's speech perception by providing detailed phonetic information about the target word. While there are several studies reporting vowel modifications in speech to infants (i.e., hyper-articulated vowels), there has been relatively little research about consonant modifications in speech to young children (except for VOT). The present study examines acoustic properties of Korean stops in Korean mothers' speech to their children aged 29 to 38 months ($N=6$). Korean tense, lax, and aspirated stops are all voiceless in word-initial position, and are perceptually differentiated by several acoustic parameters including VOT, f_0 of the following vowel, and the amplitude difference of the first and second harmonics at the voice onset of the following vowel. This study compares values of these parameters in Korean motherese to those in speech to adult Koreans from same speakers. Results focus on the acoustic properties of Korean stops in child-directed speech and how they are modified to help Korean young children learn the three-way phonetic contrast.

2pSC6. Relationship between perceived politeness and spectral characteristics of voice. Mika Ito (Dept. of Linguist., Univ. of Edinburgh, 40 George Square, Edinburgh EH8 9LL, UK, itomika@ucla.edu)

This study investigates the role of voice quality in perceiving politeness under conditions of varying relative social status among Japanese male speakers. The work focuses on four important methodological issues: experimental control of sociolinguistic aspects, eliciting natural spontaneous speech, obtaining recording quality suitable for voice quality analysis, and assessment of glottal characteristics through the use of non-invasive direct measurements of the speech spectrum. To obtain natural, unscripted utterances, the speech data were collected with a Map Task. This methodology allowed us to study the effect of manipulating relative social status among participants in the same community. We then computed the relative amplitudes of harmonics and formant peaks in spectra obtained from the Map Task recordings. Finally, an experiment was conducted to observe the alignment between acoustic measures and the perceived politeness of the voice samples. The results suggest that listeners' perceptions of politeness are determined by spectral characteristics of speakers, in particular, spectral tilts obtained by computing the difference in amplitude between the first harmonic and the third formant.

2pSC7. Size-constraints on intonation groups in speech: Evidence of an independent syllable-count principle. Annie C. Gilbert and Victor J. Boucher (Univ. of Montreal, C.P. 6128 succ. Centre-ville, Montreal, QC, Canada H3C 3J7)

This poster examines size-limits on intonation (F_0) contours in spontaneous speech and presents the results of an experiment on a syllable-count principle, which is seen to constitute, irrespective of syntax, a factor restricting the length of F_0 groups. Studies of various languages indicate a general tendency to restrict stress-groups in speech to four syllables or less. In languages where stress is not lexically coded (e.g., French), syntax is not a sufficient predictor of stress. The object was to determine whether these aspects of stress patterning also apply to tonal groups. Statistics are lacking with respect to the extent of F_0 contours in speech. Pitch-extracting software was used to analyze the speech of 15 native speakers of French (20 minutes each). Initial results suggest an eight-syllable limit on tonal groups. Based on these statistics an experiment was conducted where 40 Ss read and repeated visually presented sentences containing major syntactic divisions (phrase boundaries) at different locations. There are two central findings: (1) phrase boundaries placed at different points in

the sentence did not serve to predict tonal grouping; (2) even when the sentence structure offered the possibility of creating large tonal groups, Ss did not create contours exceeding an eight-syllable limit.

2pSC8. Sociological effects on vocal aging: Age related F_0 effects in two languages. Kyoko Nagao (Indiana Univ., Memorial Hall 406, 1021 3rd St., Bloomington, IN 47405-7005)

Listeners can estimate the age of a speaker fairly accurately from their speech (Ptacek and Sander, 1966). It is generally considered that this perception is based on physiologically determined aspects of the speech. However, the degree to which it is due to conventional sociolinguistic aspects of speech is unknown. The current study examines the degree to which fundamental frequency (F_0) changes due to advanced aging across two language groups of speakers. It also examines the degree to which the speakers associate these changes with aging in a voice disguising task. Thirty native speakers each of English and Japanese, taken from three age groups, read a target phrase embedded in a carrier sentence in their native language. Each speaker also read the sentence pretending to be 20-years younger or 20-years older than their own age. Preliminary analysis of eighteen Japanese speakers indicates that the mean and maximum F_0 values increase when the speakers pretended to be younger than when they pretended to be older. Some previous studies on age perception, however, suggested that F_0 has minor effects on listeners' age estimation. The acoustic results will also be discussed in conjunction with the results of the listeners' age estimation of the speakers.

2pSC9. Cues used for distinguishing African American and European American voices. Erik R. Thomas (Dept. of English, Box 8105, North Carolina State Univ., Raleigh, NC 27695-8105, ethomas@social.chass.ncsu.edu) and Norman J. Lass (West Virginia Univ., Morgantown, WV 26506-6122)

Past studies have shown that listeners can distinguish most African American and European American voices, but how they do so is poorly understood. Three experiments were designed to investigate this problem. Recordings of African American and European American college students performing various reading tasks were used as the basis for stimuli in all three. In the first experiment, stimuli were subjected to monotonization, lowpass filtering at 660 Hz, and no modification. In the second, stimuli featuring certain ethnically diagnostic vowels and control stimuli were subjected to monotonization, conversion of vowels to schwa, or no modification. In the third, stimuli featuring diagnostic vowels and control stimuli were modified so that the intonation of paired African American and European American speakers was swapped. In all three experiments, African American and European American listeners in North Carolina and European American listeners in West Virginia identified the ethnicity of the speaker of each stimulus. Vowel quality emerged as the most consistent cue for identifications. However, listeners accessed other cues differently for male and female speakers. Breathiness was correlated with identifications of male speakers but not of female speakers. F_0 -related factors proved more important for female speakers than for male speakers. [Work supported by NSF.]

2pSC10. Effects of aspiration on fundamental frequency in Taiwanese syllables. Yuwen Lai and Allard Jongman (Linguist. Dept., Blake Hall, Univ. of Kansas, Lawrence, KS 66044)

The perturbation effect on vowel fundamental frequency (F_0) by voiceless aspirated and unaspirated prevocalic obstruents is investigated in Taiwanese. It is well known that F_0 is significantly higher after voiceless than voiced stops. However, the perturbation effect caused by aspiration has received much less attention. Twenty-eight minimal pairs contrasting in prevocalic aspiration across three different places of articulation from seven tonal categories in Taiwanese were recorded from two male and two female speakers. An acoustic study was conducted on a total of 1120

syllables. Onset, offset, and mean F_0 as well as contours of the tones after the aspirated and unaspirated obstruents were compared. Results indicate that the onset and mean F_0 are significantly higher when following aspirated obstruents but no difference was found in offset F_0 . Tonal contour comparison shows that the F_0 raising effect triggered by aspirated obstruents disappears at 50–60% of the tone. Higher larynx and faster airflow rate are posited as the major factors for the higher F_0 after voiceless aspirated obstruents. The present results warrant a reconsideration of the traditional theory of aspiration-induced tonal split according to which a tonal split resulted from the lowering of F_0 due to aspiration.

2pSC11. Coarticulatory nasalization in modern Greek: Evidence for a link between coarticulation and syllable structure. Evanthia Diakoumakou (Oakland Univ., Linguist. Dept., Rochester, MI 48309)

An acoustic analysis of the temporal extent of vowel nasalization in the productions of six native speakers of (Standard) Modern Greek showed that the temporal extent of anticipatory vowel nasalization is limited in all contexts, although it is more extensive before tautosyllabic than heterosyllabic nasals. (On average, in stressed syllables, heterosyllabic anticipatory nasalization was 27 ms long, tautosyllabic anticipatory was 48 ms long, and carryover was 70 ms.) Modern Greek patterns in this aspect with languages like Spanish, Italian, Ikalanga, which also show a tendency toward open syllables, and obstruents are, in general, dispreferred as codas. It is hypothesized that there may be a link between a tendency for open syllables in a language and limited extent of anticipatory nasalization and it is suggested that investigation of the prosodic organization of languages may prove fruitful in determining the factors that lead to cross-language coarticulatory differences regarding vowel nasalization.

2pSC12. Examination of voicing onset time during Mandarin tone productions. Yang Chen (Dept. of Speech-Lang. Pathol., Duquesne Univ., 600 Forbes Ave., Pittsburgh, PA 15282) and Manwa Ng (Long Island Univ., Brookville, NY 11548-1300)

The voice onset time (VOT) of stop consonant is defined as the time interval between the release of a stop and the onset of the following vowel (Lisker and Abramson, 1964, 1970). Previous studies indicated that VOT could be used as a perceptual cue for identification of both voicing and place of articulation of stops (Kewley-Port, 1983; Ladefoged, 2001). Abramson (1977) suggested VOT could be interpreted as interrelated acoustic consequences of variation in the relative timing of glottal and oral gestures. This study will attempt to extend the probe on this suggestion by comparing the VOT values among different tones. Mandarin, which is a tone language, where a change in the tone of a syllable leads to a change in meaning, will be used in the study. There are four contrastive tones in Mandarin, each of which is realized by changing the vocal pitch during the course of the syllable production. The proposed study will determine if VOT values for each of the six stop sounds (/p,b,t,d,k,g/) will vary with the variation of different tones at which each of the following vowel sounds is produced. The interrelationship between VOT and tone production will be discussed.

2pSC13. A perceptual account of dissimilation. Anthony Brasher (Dept. of Linguist., Univ. of Michigan, 4080 Frieze Bldg., 105 S. State St., Ann Arbor, MI 48109-1285)

Ohala [J. J. Ohala, CLS Parasession on Language and Behavior, 178, 203 (1981)] argues that long-distance dissimilation results from hypercorrection; when confronted with words with multiple similar segments, listeners may be confused as to the origin of coarticulated features and attribute cues to one, but not all, segments. The present study evaluates this theory for a sound change involving dissimilation of breathy stops: in sequences of two or more breathy stops, all but the last become modal. If this pattern were reproduced in the laboratory then, when hearing a word with multiple breathy stops, listeners would have trouble assigning the

source of the breathy voicing. In this study, Hindi speakers heard two continua of disyllables, where the first syllable varied systematically in the temporal extent of breathy voicing. The continua differed in that the second syllable had a breathy stop in one continuum (dadhak>dhadhak) but a non-breathy stop in the other (dadak>dhadak) (dh=breathy stop). Initial results for Hindi speakers responding in a four-choice identification task support Ohala's hypothesis. Listeners identified the first syllable reliably when the second syllable is modal, but identified the first syllable inconsistently when the second syllable is breathy.

2pSC14. Pacific northwest vowels: A Seattle neighborhood dialect study. Jennifer K. Ingle, Richard Wright, and Alicia Wassink (Dept. of Linguist., Univ. of Washington, Box 354340, Seattle, WA 98195-4340, merywen@u.washington.edu)

According to current literature a large region encompassing nearly the entire west half of the U.S. belongs to one dialect region referred to as Western, which furthermore, according to Labov *et al.*, "... has developed a characteristic but not unique phonology." [http://www.ling.upenn.edu/phono_atlas/NationalMap/NationalMap.html] This paper will describe the vowel space of a set of Pacific Northwest American English speakers native to the Ballard neighborhood of Seattle, Wash. based on the acoustical analysis of high-quality Marantz CDR 300 recordings. Characteristics, such as low back merger and [u] fronting will be compared to findings by other studies. It is hoped that these recordings will contribute to a growing number of corpora of North American English dialects. All participants were born in Seattle and began their residence in Ballard between ages 0–8. They were recorded in two styles of speech: individually reading repetitions of a word list containing one token each of 10 vowels within carrier phrases, and in casual conversation for 40 min with a partner matched in age, gender, and social mobility. The goal was to create a compatible data set for comparison with current acoustic studies. F_1 and F_2 and vowel duration from LPC spectral analysis will be presented.

2pSC15. Breathiness in Indic languages. Christina Esposito, Sameeruddin Khan, and Alex Hurst (Dept. of Linguist., UCLA, 3125 Campbell Hall Box 951543, LA, CA 90095-1543, esposito@humnet.ucla.edu)

Previous work on breathiness in Indic languages has focused on the acoustic properties of breathy oral stops in languages like Hindi ([bal] hair versus [bhal] forehead) or Bengali ([baSa] house versus [bhaSa] language). However, breathiness in Indic languages often extends to nasals (e.g., Marathi ([maar] beat versus [mhaar] a caste). It is unclear if languages such as Hindi and Bengali have breathy nasals in addition to breathy oral stops. This study addresses the following questions: (1) Are breathy nasals (Nh) acoustically different from N+h sequences, both in languages where they are phonemic and ones where they are not? (2) In sequences of a breathy stop and a modal nasal (e.g., Hindi [udhmi] naughty) where is the breathiness realized, if at all? To answer these questions, audio, aerodynamic, and electroglottographic recordings will be made of Hindi, Bengali, and Marathi speakers. It is hypothesized that acoustically breathy nasals in Hindi and Bengali will not be distinct from sequences of N+ h. We believe that this will also be true for the oral stops. In addition, it is believed that in sequences of breathy oral stop followed by a modal nasal (e.g., ChN), the breathiness will be produced on the nasal.

2pSC16. Low vowels and transparency in Kinande vowel harmony. Bryan Gick, Douglas Pulleyblank (Dept. of Linguist., Univ. of British Columbia, E270-1866 Main Mall, Vancouver, BC, Canada V6T 1Z1, gick@interchange.ubc.ca), Nguessimo Mutaka (Univ. of Yaounde 1, Cameroon), and Fiona Campbell (Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z1)

Transparency—in which a harmony effect passes over a segment without affecting it phonetically or phonologically—has been a controversial concept in previous literature on harmony systems. A typical case of so-called transparency involves cross-height vowel harmony in Kinande, a Bantu language (J.40). Previous accounts have analyzed low vowels in this system as being transparent to harmony [Schlindwein, *NELS* 17, 551–567 (1987)]. Further, some analysts have considered low vowels theoretically incapable of undergoing tongue root harmony. These claims were tested in a single-subject field study using ultrasound imaging to measure tongue root position in low vowels. Results indicate that (a) advanced versus retracted tongue root position (ATR) is a viable feature for describing the phonological distinction in the vowel system; (b) there is a phonetic difference between low vowels when adjacent to ATR triggering vowels; (c) this distinction in low vowels does not decrease with distance from trigger vowels, suggesting that these vowels are undergoing phonological harmony rather than phonetic assimilation; and finally, (d) the ATR distinction is phonetically categorical in high vowels, but shows crossover in mid and low vowels. Implications for phonological theory and phonetics-phonology interface will be discussed. [Work supported by NSERC and SSHRC.]

2pSC17. The role of psychophysical difference in the discrimination of non-native contrasts. James Harnsberger, Rahul Shrivastav (Dept. Comm. Sci. and Disord., Univ. of Florida, Gainesville, FL 32611), and Mark Skowronski (Univ. of Florida, Gainesville, FL 32611)

Models of cross-language speech perception have shown only limited success in predicting the discriminability of non-native contrasts. These failures may be attributed partly to an inability to quantify the phonetic similarities between non-native speech sounds and between non-native speech sounds and native speech categories. This study represents an attempt to quantify gross psychophysical differences between consonants (e.g., the absolute difference between two consonants as represented in the peripheral auditory system). Two metrics were evaluated with a set of nasal consonant place contrasts from Malayalam. The first focused only on formant transitions at nasal-vowel boundaries (critical cue measure) while the second compared the spectrum of temporal windows spanning the entire length of both syllables that constituted a contrast (whole stimulus measure). The critical cue measure was presumed to have the advantage of incorporating only perceptually relevant information, while the whole stimulus measure was thought to be preferable if the relevant cues were poorly understood and/or were distributed throughout the syllable. The results showed a significant correlation ($r=0.63$) between the critical cue measure and discrimination scores. The whole stimulus measure showed only a weak correlation ($r=0.35$) unless a set of outliers (bilabial contrasts) was excluded ($r=0.80$).

2pSC18. Comparing listener preferences for text-level prosody in English and French. Caroline Smith (Dept. of Linguist., Univ. of New Mexico, MSC 03-2130, Albuquerque, NM 87131-1196)

Cross-language prosodic differences are well-known at the phrasal level and below, and a few studies have shown language-dependent differences in speakers' prosody over a discourse or text [Fon (2002), Smith and Hogan (2003)]. For comparison with the patterns found in production, this study examines listeners' preferences in text-level prosody. Previously-analyzed recordings of one speaker of American English reading aloud an English text, and one speaker of Parisian French reading a comparable French text, were used as a basis. The discourse organization of each text was analyzed by several native speakers who categorized each sentence-to-sentence transition as a Topic Shift, Continuation or Elaboration.

The original recordings were manipulated in several ways to alter rate, sentence-final lengthening and pause duration, and the manipulated versions were presented to listeners. English listeners liked best those versions where values for rate and lengthening matched the speakers' means for each separate type of topic transition. French listeners preferred the version where each prosodic variable kept the same mean value throughout. These results are supported by acoustic measurements which also point to a more salient role for rhythmic effects at the discourse level in English than in French. [Work supported by NSF grant BCS-9983106.]

2pSC19. Identifying native Cantonese stops: Implication for native speakers perception. Man Gao (Dept. of Linguist., Yale Univ., New Haven, CT 06511)

A series of perception tasks were conducted to probe Cantonese speakers' ability to perceive their languages syllable-final unreleased stops. Stimuli consisted of minimal sets of monosyllabic words with contrasting codas produced by a speaker of Hong Kong Cantonese. Twelve Guangzhou Cantonese speakers and five Hong Kong Cantonese speakers performed the perceptual tasks which included: matching stimuli with the corresponding orthographic character; AXB discrimination; and identifying word-final [p], [t], [k] and vowel explicitly. Five native speakers of Mandarin (where final stops are not allowed) also took the AXB and identification tasks. Both Guangzhou and Hong Kong speakers scored high in matching task: 93 and 96 percent correct respectively; for AXB task, subjects from Hong Kong (93%) did slightly better than those from Guangzhou (85%), and Mandarin speakers scored the lowest (78%). The most interesting findings were in the identification tasks, although subjects from Hong Kong scored the highest (82%), the accuracy of Guangzhou Cantonese speakers was surprisingly low (49%)—even worse than Mandarin speakers (69%). Explanation for the observed discrepancy between Guangzhou and Hong Kong subjects will be discussed with respect to their different language education as well as their distinct language experience. Implications for phonological awareness will also be discussed.

2pSC20. Processing voiceless vowels in Japanese: Effects of language-specific phonological knowledge. Naomi Ogasawara (Dept. of Linguist., Univ. of Arizona, Douglass 200E, P.O. Box 210028, Tucson, AZ 85721)

There has been little research on processing allophonic variation in the field of psycholinguistics. This study focuses on processing the voiced/voiceless allophonic alternation of high vowels in Japanese. Three perception experiments were conducted to explore how listeners parse out vowels with the voicing alternation from other segments in the speech stream and how the different voicing statuses of the vowel affect listeners' word recognition process. The results from the three experiments show that listeners use phonological knowledge of their native language for phoneme processing and for word recognition. However, interactions of the phonological and acoustic effects are observed to be different in each process. The facilitatory phonological effect and the inhibitory acoustic effect cancel out one another in phoneme processing; while in word recognition, the facilitatory phonological effect overrides the inhibitory acoustic effect.

2pSC21. The influence of stress on some acoustic correlates to the stop voicing distinction in French. Nassima Abdelli-Beruh (Dept. of Speech-Lang. Pathol. and Audiol., New York Univ., 719 Broadway Ste. 200, New York, NY 10003 and Dept. of Speech and Hearing Sci., City Univ. of New York, 365 Fifth Ave., New York, NY 10016) and Radhika Aravamudhan (Post-Doc Res. Fellow at Boys Town Natl. Res. Hospital, Omaha, NE)

This study examined how monolingual French speakers produced the stop voicing distinction in stressed and unstressed syllable-initial stops. Syllables were embedded in sentences. Voicing-related differences in du-

rations of VOT, closure and vowel were calculated and analyzed as a function of stress (stressed on the target syllable, stress on the syllable preceding the target syllable). Percentages of closures with voicing were tallied as function of the voicing category of the stops and the stressed condition. Results from ANOVA showed that the absolute durations were smaller in the unstressed than in the stressed condition. The magnitude of the voicing-conditioned related duration differences in VOT, closure and vowel were also influenced by stress.

2pSC22. The effects of implosives and prenasalized stops on pitch in Shona. Mario E. Chavez-Peon (Dept. of Linguist., Univ. of British Columbia, E-270, 1866 Main Mall, Vancouver, BC, Canada V6T 1Z1)

It is well known that F_0 at vowel onset can be influenced by a preceding consonant. That influence varies significantly across languages and consonant types, and may function as a perceptual signal to consonant manner. It has further been suggested that tone languages may behave differently from non-tone languages in this respect, with a shorter duration of consonantal perturbation [Hombert, *Studies in African Linguistics*, 1977]. Previous studies include a limited range of consonant types, and too few tone languages to test Hombert's proposal. This study presents the results of an acoustical investigation of the effects of implosives and prenasalized stops on the F_0 of a following vowel in Shona, a tone language. It is found that implosives have a similar raising effect on F_0 at vowel onset than that of voiceless (aspirated) stops, contrary to expectations based on previous studies [Wright and Shryock, *Journal of the Phonetic Association*, 1993]. It is also found that prenasalized consonants behave as nasals, having no effect on the F_0 of the following vowel, again contrary to expectation [cf. Trithart, *Studies in Bantu Tonology*, 1976; and Hombert, *Studies in Bantu Tonology*, 1976]. Finally, duration results do not support Hombert's position regarding tone languages.

2pSC23. Revisiting the Canadian English vowel space. Robert Hagiwara (Linguist. Dept., Univ. of Manitoba, Winnipeg, MB, Canada R3T 5V5)

In order to fill a need for experimental-acoustic baseline measurements of Canadian English vowels, a database is currently being constructed in Winnipeg, Manitoba. The database derives from multiple repetitions of fifteen English vowels (eleven standard monophthongs, syllabic /r/ and

three standard diphthongs) in /hVd/ and /hVt/ contexts, as spoken by multiple speakers. Frequencies of the first four formants are taken from three timepoints in every vowel token (25, 50, and 75% of vowel duration). Preliminary results (from five men and five women) confirm some features characteristic of Canadian English, but call others into question. For instance the merger of low back vowels appears to be complete for these speakers, but the result is a lower-mid and probably rounded vowel rather than the low back unrounded vowel often described. With these data Canadian Raising can be quantified as an average 200 Hz or 1.5 Bark downward shift in the frequency of F_1 before voiceless /t/. Analysis of the database will lead to a more accurate picture of the Canadian English vowel system, as well as provide a practical and up-to-date point of reference for further phonetic and sociophonetic comparisons.

2pSC24. Concurrent recognition of focus and question in Mandarin. Fang Liu (Univ. of Chicago, Dept. of Linguist., 1010 E. 59th St., Chicago, IL 60637)

F_0 has been shown to vary not only with lexical tones but also with focus and sentence type (statement versus question) in Mandarin. This study further investigates whether listeners can detect concurrent focus and sentence type information in speech. 320 statements and unmarked questions with initial, medial, final, or neutral focus produced by two speakers were tested on 11 listeners. Results show that listeners can identify both sentence type and focus in most cases (89.1% and 88.7%, respectively), suggesting that F_0 variations related to the two functions can be simultaneously perceived. Meanwhile, the lowest rates were found for identifying neutral focus in questions (71%) and for identifying statement in sentences with final focus (78%). In both cases the confusions seem to arise from the conflicting F_0 adjustments by sentence type and focus in the sentence-final position (final F_0 being raised for both question and final focus, but not for statement). Overall, the findings not only indicate that F_0 can concurrently encode both focus and sentence type in addition to tonal information, but also demonstrate how delicate a balance is maintained between these communicative functions when they have to share the same articulatory/acoustic parameter for their respective manifestations.

2p TUE. PM

TUESDAY AFTERNOON, 17 MAY 2005

GEORGIA B, 1:30 TO 5:00 P.M.

Session 2pUW

Underwater Acoustics: Effects of Environmental Uncertainty

David R. Dowling, Chair

Dept. of Mechanical Engineering, Univ. of Michigan, 1231 Beal Ave., Ann Arbor, MI 48109-2133

Contributed Papers

1:30

2pUW1. Uncertainty propagation in ocean acoustic waveguides. Steven Finette (Acoust. Div., Naval Res. Lab., Washington, DC 20375)

This presentation discusses a general method for representing uncertainty in numerical models that describe acoustic field propagation within partially specified ocean waveguide environments. The approach treats uncertainty within a probabilistic framework, representing both the sound speed distribution and acoustic field as stochastic processes. In this talk, environmental uncertainty is linked to incomplete knowledge of the vol-

ume sound speed field and is directly incorporated into the mathematical formulation for acoustic propagation in a stochastic ocean waveguide. As an example, equations describing the propagation of uncertainty [Finette, *J. Acoust. Soc. Am* (in press)] are solved analytically for the special case of a point source in a random, but spatially uniform sound speed field bounded by a pressure release surface and a rigid bottom. Results for the first two moments of the acoustic field are calculated using this approach and compared with an independent estimate of the same moments obtained from Monte Carlo based realizations that are computed from a deterministic parabolic wave equation. [Work supported by ONR.]

2pUW2. Converting sound speed uncertainty to an effective correlated noise term in array processing. Jit Sarkar, Bruce Cornuelle, Philippe Roux, W. S. Hodgkiss, W. A. Kuperman (Scripps Inst. of Oceanogr., UCSD, 9500 Gilman Dr., Mail Code 0238, La Jolla, CA 92093-0238, jit@mpl.ucsd.edu), and Mark Stevenson (NATO SACLANT Undersea Res. Ctr., La Spezia 19138, Italy)

The sensitivity of the received acoustic field at an array from a specific source to perturbations of the sound speed field about a mean value can be computed using a Born approximation formulation involving forward model calculations. When this sensitivity map is appropriately combined with a covariance of the uncertainty of the sound speed perturbations, one obtains the resultant perturbation to the cross spectral density matrix (CSDM) of the acoustic field. This additional term to the mean CSDM can be thought of as a noise term resulting from the uncertainty in the sound speed field. Computations indicate that this additional term is a form of correlated noise that suggests the possibility of mitigating environmental uncertainty with a robust array processing procedure.

2:00

2pUW3. Rapid acoustic field computations within environmental uncertainty bounds. Kevin D. Heaney (OASIS, Inc., Falls Church, VA 22044)

Oceanographic and geo-acoustic variability in the shallow water can dominate acoustic propagation. Often the environment is critically under-sampled, compared to the sensitivity of the full-field propagation model used [such as the Parabolic Equation (PE) model]. For regions with environmental uncertainty (which is known from measurements or models), there will be a Minimum Achievable Uncertainty (MAU) in acoustic propagation modeling (Transmission Loss). Given this MAU, it is proposed that field interpolation techniques be applied to speed up the acoustic propagation computation, consistent with keeping the interpolation errors below the MAU. Techniques for estimating the uncertainty in TL and speeding up the broadband range-dependent computation using the Navy Standard PE model will be presented. For fully 3-D field computations in a range-dependent environment, the speed up is expected to be several orders of magnitude.

2:15

2pUW4. In-situ estimates of environmental uncertainty on sonar performance prediction. Kevin D. Heaney (OASIS, Inc., Falls Church, VA 22044)

Oceanographic and geo-acoustic variability in the shallow water can dominate acoustic propagation. Sonar performance predictions using archival sound speed and geo-acoustic information do not contain any information about variability or uncertainty in the environment and its likely impact on sonar performance prediction. Using measurements from the NATO Undersea Research Centres BOUNDARY 2003 sea test, an approach is presented that estimates the uncertainty and transfers it through the acoustic propagation and the sonar equation to sonar performance. The sound speed field is characterized by an Empirical Orthogonal Function decomposition of the measured profiles. The geo-acoustic uncertainty is considered to be a normal distribution around the inverted geo-acoustic parameters (from 4 separate inversions). Dynamic ambient noise variability is measured from the array beamformer output. Uncertainty in propagation is quantified in terms of sound speed, geo-acoustic, and ambient noise uncertainty. The uncertainty in sonar performance, due to environmental uncertainty, is then displayed to the operator. [Work supported by the ONR Capturing Uncertainty DRI.]

2pUW5. Performance of mode-based processing in the presence of environmental uncertainty. Lisa Zurk (Elec. & Comput. Eng., Portland State Univ., P.O. 751, Portland, OR 97207)

A critical limiting factor to the performance of many underwater sonar detection and localization techniques is the typically high degree of uncertainty regarding the ocean environment. Several researchers have proposed mode-based processing methods for Matched Field Processing, depth-dependent pre-filtering, and sub-array processing. Provided that an appropriate modal decomposition can be applied, these methods are believed to provide processing that is less sensitive to environmental mismatch, with the added potential advantage of computational efficiency. In this paper, we attempt to quantify the performance of mode-based processing in uncertain environments, and compare this performance to more classic processing approaches. Results are presented from Monte Carlo simulations of range-dependent environments in which the uncertainty is varied parametrically. An analytical framework for describing the uncertainty is presented and discussed.

2:45

2pUW6. A high frequency time reversal array deployment in a very shallow water environment. Thomas Folegot (NATO Undersea Res. Ctr., MILOC, Viale San Bartolomeo 400, 19138 La Spezia, Italy), Dominique Clorennec, Jean-Gabriel Minonzio, Julien de Rosny, Claire Prada (Laboratoire Ondes et Acoustique, ESPCI, 75005 Paris, France), Mathias Fink (Laboratoire Ondes et Acoustique, ESPCI, 75005 Paris, France), Sidonie Hibrat, and Lothar Berniere (Atlantide, Technopole Brest Iroise, 29238 Brest Cedex 3, France)

A new rigid 24-element acoustic source receiver array in the 10 to 15 kHz frequency band has recently been developed [Folegot *et al.*, submitted to OCEANS2005-Europe]. Despite of a height of 10 m, it has been specified to be easily deployed from a pier and to allow low cost at-sea experiments with water depth up to 12 m. The array has been deployed in a pool of 50 m length and 10 m depth, where calibrated surface plane waves have been generated. The objective of the trial was to undertake a first deployment of the system, to calibrate the instrument and provide some scientific results based on time reversal in a very controlled experimental setup. The Time Reversal Operator has been measured in several configurations. Subsequently, its eigenvectors associated with target reflections have been emitted by the mirror in order to focus separately on each target. The sensitivity of the method against target depth, target type, emission basis, and surface roughness has been investigated. The collected data will help to improve the understanding of environmental factors affecting acoustic propagation in the 10–15 kHz range and will make an important contribution to the development of advanced acoustic communications methods and tomographic capabilities.

3:00–3:15 Break

3:15

2pUW7. Improvement of time reversal communications in underwater acoustic channels using adaptive equalizers. Heechun Song, William S. Hodgkiss, William A. Kuperman, Philippe Roux (MPL/SIO, UCSD, La Jolla, CA 92093), Tuncay Akal (MPL/SIO, UCSD, La Jolla, CA 92093), and Mark Stevenson (NATO Undersea Res. Ctr., La Spezia, Italy)

The spatial and temporal focusing properties of time reversal methods can be exploited for undersea acoustic communications. Spatial focusing mitigates channel fading and produces a high signal-to-noise ratio at the intended receivers along with a low probability of interception elsewhere. While the temporal focusing (compression) reduces significantly intersymbol interference (ISI), there always is some residual ISI depending upon the number of transmitters and the complexity of the channel. Moreover, a slight change in the environment over the two-way propagation interval introduces additional ISI. Using shallow water experimental data, we dem-

onstrate that the performance of multiple-input/multiple-output time reversal communications can be improved significantly by cascading the received time series with adaptive equalizers to remove the residual ISI.

3:30

2pUW8. Broadband source localization using matched correlation processing. Gordon R. Ebbeson, Marie-Noël R. Matthews, Garry J. Heard, Francine Desharnais (DRDC Atlantic, P.O. Box 1012, 9 Grove St., Dartmouth, NS, Canada B2Y 3Z7), and David J. Thomson (Victoria, BC, Canada V9C 4A4)

For many years, model-based signal processing algorithms using Matched Field Processing (MFP) techniques have been analyzed with the goal of improving the capability of passive sonar systems for localizing quiet underwater sources. Recently, researchers at DRDC Atlantic have been investigating Matched Correlation Processing (MCP) as a faster alternative to MFP. In this method, the cross-correlations for a source as measured with a pair of hydrophones in a horizontal array are matched with those generated with a correlation model for many candidate ranges and depths along a candidate bearing. These matches are carried out with a number of hydrophone pairs to form many ambiguity surfaces. The maximum on the average of these surfaces is assumed to yield the best estimate of the source position. By carrying out this procedure over a number of candidate bearings, a full 3-D search for the source location is achieved. Since 2002, a number of localization trials have been carried out east of Nova Scotia, Canada. During those trials, an array was deployed on the sea floor and used to collect acoustic signals from various broadband sources. In this paper, we describe the broadband MCP localization technique and show some localization results from those trials.

3:45

2pUW9. Broadband continuous wave source localization via pair-wise, cochleagram processing. Eva-Marie Nosal and L. Neil Frazer (Univ. of Hawaii at Manoa, 1680 East-West Rd., Post 813, Honolulu, HI 96822)

A pair-wise processor has been developed for the passive localization of broadband continuous-wave underwater sources. The algorithm uses sparse hydrophone arrays and does not require previous knowledge of the source signature. It is applicable in multiple source situations. A spectrogram/cochleagram version of the algorithm has been developed in order to utilize higher frequencies at longer ranges where signal incoherence, and limited computational resources, preclude the use of full waveforms. Simulations demonstrating the robustness of the algorithm with respect to noise and environmental mismatch will be presented, together with initial results from the analysis of humpback whale song recorded at the Pacific Missile Range Facility off Kauai. [Work supported by MHPCC and ONR.]

4:00

2pUW10. Modeling of forward sector adaptive matched field processing. Jeffrey Dunne (Johns Hopkins Univ. Appl. Phys. Lab., 11100 Johns Hopkins Rd., Laurel, MD 20723, Jeffrey.Dunne@jhuapl.edu)

A modeling study was undertaken to examine the potential benefit of adaptive matched field processing (AMFP) to the forward sector capability of single-line, twin-line, and volumetric arrays. Comparisons are made with conventional MFP (CMFP) and plane-wave beamforming (PWB) in order to assess the degree of ownship noise reduction obtainable and any corresponding improvement to the signal-to-noise ratio (SNR). A minimum variance distortionless response beamformer using dominant mode rejection was implemented, applied to both uniform and distorted array shapes. Significant improvement over CMFP and PWB in tracking and SNR was seen for modeled data in both cases, with the distorted array showing, not surprisingly, better left-right rejection capability. A brief investigation of the effects of environmental mismatch was undertaken in order to assess its potential impact on the value-added of AMFP. [Work was supported by the Defense Advanced Research Projects Agency, Advanced Technology Office.]

2pUW11. An iso-deviant approach for acoustic computations using efficient adaptive gridded for littoral environments. Erik R. Rike and Donald R. DelBalzo (Neptune Sci. Div. of Planning Systems, Inc., 40201 Hwy 190 E, Slidell, LA 70461, delbalzo@neptunesci.com)

Transmission Loss (TL) computations in littoral areas require a dense spatial and azimuthal grid to achieve acceptable accuracy and detail. The computational cost of accurate predictions led to a new concept, OGRES (Objective Grid/Radials using Environmentally-sensitive Selection), which produces sparse, irregular acoustic grids, with controlled accuracy. Recent work to further increase accuracy and efficiency with better metrics and interpolation led to EAGLE (Efficient Adaptive Gridded for Littoral Environments). On each iteration, EAGLE produces grids with approximately constant spatial uncertainty (hence, iso-deviance), yielding predictions with ever-increasing resolution and accuracy. The EAGLE point-selection mechanism is tested using the predictive error metric and 1-D synthetic data-sets created from combinations of simple signal functions (e.g., polynomials, sines, cosines, exponentials), along with white and chromatic noise. The speed, efficiency, fidelity, and iso-deviance of EAGLE are determined for each combination of signal, noise, and interpolator. The results show significant efficiency enhancements compared to uniform grids of the same accuracy. [Work sponsored by ONR under the LADC project.]

4:30

2pUW12. Adapting a truly nonlinear filter to the ocean acoustic inverse problem. Andrew A. Ganse and Robert I. Odom (Appl. Phys. Lab. (APL-UW), 1013 NE 40th St., Seattle, WA 98105, aganse@apl.washington.edu)

Nonlinear inverse problems including the ocean acoustic problem have been solved by Monte Carlo, locally-linear, and filter based techniques such as the Extended Kalman Filter (EKF). While these techniques do provide statistical information about the solution (e.g., mean and variance), each suffers from inherent limitations in their approach to nonlinear problems. Monte Carlo techniques are expensive to compute and do not contribute to intuitive interpretation of a problem, and locally-linear techniques (including the EKF) are limited by the multimodal objective landscape of nonlinear problems. A truly nonlinear filter, based on recent work in nonlinear tracking, estimates state information for a nonlinear problem in continual measurement updates and is adapted to solving nonlinear inverse problems. Additional terms derived from the system's state PDF are added to the mean and covariance of the solution to address the nonlinearities of the problem, and overall the technique offers improved performance in nonlinear inversion. [Work supported by ONR.]

4:45

2pUW13. A neural network for mode space source depth estimation. William Lee and Yung Lee (Sci. Applications Intl. Corp., 1710 SAIC Dr., McLean, VA 22102)

A great deal of research in mode space source localization processing exists. One technique proposed by Shang performs mode filtering of the observed pressure field on a vertical line array, using the calculated mode amplitude distributions and phase differences to estimate source location. Ozard *et al.* proposed a feed forward neural network using input vectors of mode energy distribution for source localization, a dedicated neuron for each possible source location was employed in their study. Instead of employing a dedicated neuron for each possible search location, the interest of this study is to determine an efficient setup of neurons to properly sample the search space. In MATLAB neural network toolbox, the self-organizing-map function learns to classify input vectors according to how they are grouped in the input space, specifically their distribution and topology. We used this function to examine how input vectors of mode energy distribution for signals at different depths are grouped to determine the sampling grids of the search space in depth. We also compared the neural network performance using input vectors of mode amplitude distribution with those using input vectors of mode energy distribution.

Meeting of Accredited Standards Committee (ASC) S1 Acoustics

J. P. Seiler, Chair S1

*U.S. Department of Labor, Mine Safety and Health Admin., P.O. Box 18233, Bldg. 38,
Cochrans Mill Road, Pittsburgh, PA 15236*

G. S. K. Wong, Vice Chair S1

*Institute for National Measurement Standards, National Research Council, Montreal Road, Bldg. M36,
Ottawa, Ontario K1A 0R6, Canada*

Accredited Standards Committee S1 on Acoustics. Working group chairs will report on the status of standards currently under development in the areas of physical acoustics, electroacoustics, sonics, ultrasonics, and underwater sound, etc. Consideration will be given to new standards that might be needed over the next few years. Open discussion of committee reports is encouraged.

People interested in attending the meeting of the TAGs for ISO/TC 43 Acoustics and IEC/TC 29 Electroacoustics, take note - those meetings will be held in conjunction with the Standards Plenary meeting at 10:30 a.m. on Tuesday, 17 May 2005.

Scope of S1: Standards, specifications, methods of measurement and test, and terminology in the field of physical acoustics, including architectural acoustics, electroacoustics, sonics and ultrasonics, and underwater sound, but excluding those aspects which pertain to biological safety, tolerance and comfort.

Meeting of Accredited Standards Committee (ASC) S12 Noise

R. D. Hellweg, Chair S12

Hewlett Packard Co., Acoustics Lab, MR01-1/P18, 200 Forest Street, Marlborough, MA 01752

R. D. Godfrey, Vice Chair S12

Owens Corning Science & Technology Center, 2790 Columbus Road, Route 16, Granville, OH 43023-1200

Accredited Standards Committee S12 on Noise. Working group chairs will report on the status of noise standards currently under development. Consideration will be given to new standards that might be needed over the next few years. Open discussion of committee reports is encouraged.

People interested in attending the meeting of the TAG for ISO/TC 43/SC 1 Noise, take note - that meeting will be held in conjunction with the Standards Plenary meeting at 10:30 a.m. on Tuesday, 17 May 2005.

Scope of S12: Standards, specifications, and terminology in the field of acoustical noise pertaining to methods of measurement, evaluation and control, including biological safety, tolerance and comfort and physical acoustics as related to environmental and occupational noise.

Session 3aAA

**Architectural Acoustics, Speech Communication and Psychological and Physiological Acoustics:
Possibilities and Problems with Auralizations of Classroom Acoustics**

Robert C. Coffeen, Chair

*Univ. of Kansas, School of Architectural and Urban Design, Lawrence, KS 66045***Chair's Introduction—8:30***Invited Papers***8:35****3aAA1. Predicting and auralizing acoustics in classrooms.** Claus Lyng Christensen (Odeon A/S c/o Acoust. Technol., Oersted DTU, Bldg. 352, DK-2800 Lyngby, Denmark)

Although classrooms have fairly simple geometries, this type of room is known to cause problems when trying to predict their acoustics using room acoustics computer modeling. Some typical features from a room acoustics point of view are: Parallel walls, low ceilings (the rooms are flat), uneven distribution of absorption, and most of the floor being covered with furniture which at long distances act as scattering elements, and at short distance provide strong specular components. The importance of diffraction and scattering is illustrated in numbers and by means of auralization, using ODEON 8 Beta.

8:55**3aAA2. Subjective evaluation of auralizations created from multi-channel anechoic recordings of a talker in motion.** Michelle C. Vigeant and Lily M. Wang (Architectural Engr. Prog., Univ. of Nebraska–Lincoln, Peter Kiewit Inst., 1110 S. 67th St., Omaha, NE 68182-0681, mvigeant@unlnotes.unl.edu)

A high degree of speech intelligibility is very important in educational environments. When designing such spaces, like classrooms, auralizations can be used to subjectively assess the degree of speech intelligibility and clarity. Auralizations are most commonly made by convolving the impulse response (IR) of an omni-directional source with a single channel anechoic speech recording. This paper explores the idea of using multi-channel recordings to create the auralizations, using a female talker in motion. An omni-directional source is split into quadrants and the IR is calculated for each section. These IR's are convolved with the appropriate channel of the anechoic recording and then the four auralizations are mixed to create one final auralization. The auralizations were made using four-channel anechoic recordings of a person walking on a platform while talking. Subjective tests were conducted to determine the ease with which subjects could identify the direction of the movement of the source in rooms with varying amounts of absorption. This method can be used to create more realistic classroom auralizations, as teachers typically move around the room as they teach. [Work supported by the National Science Foundation.]

9:15**3aAA3. ClassTalk system for predicting and auralizing speech in noise with reverberation in classrooms.** Murray Hodgson and Daniel Graves (UBC Acoust. & Noise Res. Group, SOEH, 3rd Fl., 2206 East Mall, Vancouver, BC, Canada V6T 1Z3)

This paper discusses and demonstrates the ClassTalk system for predicting, visualizing and auralizing speech in noise with reverberation in classrooms. The classroom can contain a speech-reinforcement system (SRS). Male or female speech sources, SRS loudspeakers and overhead, slide or digital projectors, or ventilation-noise sources, can have four output levels. Empirical models are used to predict speech and noise levels, and Early Decay Times, from which Speech Transmission Index (STI) and Speech Intelligibility (SI) are calculated. ClassTalk visualizes the floor-plan, speech- and noise-source positions, and the receiver position. The user can walk through the room at will. In real time five quantities, background-noise level, speech level, signal-to-noise difference, STI and SI, are displayed along with occupied and unoccupied reverberation times. The sound module auralizes male or female speech mixed with the relevant noise signals, with predicted, frequency-varying reverberation superimposed using MaxxVerb. Technical issues related to the development of the sound module are discussed. The potential of the systems auralization module for demonstrating the effects of the acoustical environment and its control on speech is discussed and demonstrated.

9:35**3aAA4. Comparison of speech intelligibility scores for direct listening and headphone playback.** John S. Bradley, Hiroshi Sato, Bradford N. Gover, and Natalie York (Inst. for Res. in Construction, Natl. Res. Council, Montreal Rd., Ottawa, Canada K1A 0R6)

Speech recognition test scores from direct listening and headphone playback were compared to validate the use of headphone playback for investigating the effects of varied room acoustics conditions. Varied acoustical conditions for direct listening, that were representative of classrooms, were achieved using simulated sound fields in an anechoic room with an 8-channel electro-acoustic simulation system. Eight conditions consisting of the combinations of 2 reverberation times and 4 signal-to-noise ratios were used. By recording these conditions with an acoustical mannequin, headphone playback of the same conditions could be achieved and the corresponding speech recognition scores compared. Two different approaches for correcting for the microphone and headphone

responses were compared. Subjects also responded to a short questionnaire to describe the audible differences between direct playback and headphone listening. Because spatial characteristics were frequently found to differ between the two listening modes, further comparisons were carried out using sound fields with systematic differences in the locations of the speech and noise sources. Although there were always some audible differences between the two playback modes, it was possible to obtain the same speech recognition scores using headphone playback as with direct listening. [Work supported by CLLRnet.]

9:55

3aAA5. Auralizations and other computer model studies to predict qualitative and quantitative measures of speech intelligibility in classrooms. Gary Siebein, Hyeong-seok Kim (Univ. of Florida School of Architecture, P.O. Box 115702, Gainesville, FL 32607), and Hyun Paek (Siebein Assoc., Inc., Consultants in Architectural Acoust., Gainesville, FL 32607)

Auralizations and other computer model studies were used to predict qualitative and quantitative measures of speech intelligibility in classrooms under realistic conditions of background noise and reverberation. Speech intelligibility tests were given to college students in two classrooms and one racquetball court at 5 signal-to-noise ratios. Auralizations of the speech intelligibility tests were made from computer. Speech intelligibility tests were then administered in a sound booth using the auralized material. Fifteen different acoustical measurements related to speech intelligibility were also made at multiple locations in the actual classrooms and in the computer models of the classrooms. The scores on the speech intelligibility tests given in the actual rooms in the five noise conditions were closely duplicated in the equivalent tests conducted in a sound booth using the simulated speech signals obtained in the computer models. Both quantitative and qualitative measures of speech intelligibility in the actual rooms were accurately predicted in the computer models. Correlations (R2) between acoustical measures made in the full size classrooms and the computer models of the classrooms of 0.92 to 0.99 were found.

10:15–10:30 Break

10:30

3aAA6. Classroom auralizations using both speech and intruding noise. Robert C. Coffeen (School of Architecture and Urban Design, Marvin Hall, The Univ. of Kansas, Lawrence, KS 66045)

Continuing work with computer auralizations for education spaces—Can realistic speech auralizations be obtained for classrooms and similar spaces with disturbing ambient noise as produced by HVAC systems, noise from adjacent spaces, aircraft flyovers, and other common noise sources? Several auralizations will be presented relating to these situations.

10:50

3aAA7. Classroom sound quality—a comparison of recordings to virtual auralization. Kenneth P. Roy and Sean D. Browne (Armstrong Innovation Ctr., 2500 Columbia Ave., Lancaster, PA 17604)

Sound quality in classrooms is an important consequence of the architectural and mechanical systems design. And it is the primary factor in determining the potential for speech intelligibility, which of course is also dependent on individual listening characteristic and capabilities. So the question is “how do we assess the potential for good sound quality within the design phase of school construction?” An obvious solution is to use architectural modeling and auralization . . . but the question here is “how good are these virtual techniques?” as with the use of EASE software. We will present 1st a sample of the modeling and auralization process showing the capabilities via the generated sound files. Next we will listen to recordings made in an occupied classroom both before and after architectural interventions. Then we will listen to a series of auralized classroom settings. And finally, we will compare recordings to auralizations for the same classroom. One important aspect of “real room” sound quality is occupant generated noise (non-speech), which is not an easy aspect to include in virtual models.

11:10

3aAA8. Rapid Speech Transmission Index predictions and auralizations of unusual instructional spaces at MIT’s new Stata Center. David A. Conant (McKay Conant Brook Inc., 5655 Lindero Canyon Rd., Westlake Village, CA 91326)

The Stata Center for Computer, Information and Intelligence Sciences, recently opened at the Massachusetts Institute of Technology, includes a variety of oddly-shaped seminar rooms in addition to lecture spaces of somewhat more conventional form. The architects design approach prohibited following conventional, well understood room-acoustical behavior yet MIT and the design team were keenly interested in ensuring that these spaces functioned exceptionally well, acoustically. CATT-Acoustic room modeling was employed to assess RASTI through multiple design iterations for all these spaces. Presented here are computational and descriptive results achieved for these rooms which are highly-regarded by faculty. They all sound peculiarly good, given their unusual form. In addition, binaural auralizations for selected spaces are provided.

Session 3aABa

Animal Bioacoustics: General Topics in Hearing and Auditory Systems

Whitlow W. L. Au, Chair

Hawaii Inst. of Marine Biology, P.O. Box 1106, Kailua, HI 96734

Chair's Introduction—7:55

Contributed Papers

8:00

3aABa1. Hearing and the round goby: Understanding the auditory system of the round goby (*Neogobius melanostomus*). Andrea J. Belanger and Dennis M. Higgs (Dept. of Biological Sci., Univ. of Windsor ON, 401 Sunset Ave., Windsor, ON, N9B 3P4, Canada)

The round goby (*Neogobius melanostomus*), is an invasive species in the Great Lakes watershed. Adult round gobies show behavioral responses to conspecific vocalizations but physiological investigations have not yet been conducted to quantify their hearing abilities. We have been examining the physiological and morphological development of the auditory system in the round goby. Various frequencies (100 Hz to 800 Hz and conspecific sounds), at various intensities (120 dB to 170 dB re 1 Pa) were presented to juveniles and adults and their auditory brain-stem responses (ABR) were recorded. Round gobies only respond physiologically to tones from 100–600 Hz, with threshold varying between 145 to 155 dB re 1 Pa. The response threshold to conspecific sounds was 140 dB re 1 Pa. There was no significant difference in auditory threshold between sizes of fish for either tones or conspecific sounds. Sacculus epithelia were stained using phalloidin and there was a trend towards an increase in both hair cell number and density with an increase in fish size. These results represent a first attempt to quantify auditory abilities in this invasive species. This is an important step in understanding their reproductive physiology, which could potentially aid in their population control. [Funded by NSERC.]

8:15

3aABa2. Sound localization and auditory response capabilities in round goby (*Neogobius melanostomus*). Audrey K. Rollo and Dennis M. Higgs (Dept. of Biol., Univ. of Windsor, Windsor, ON, Canada N9B 3P4, audrey-rollo@hotmail.com)

A fundamental role in vertebrate auditory systems is determining the direction of a sound source. While fish show directional responses to sound, sound localization remains in dispute. The species used in the current study, *Neogobius melanostomus* (round goby) uses sound in reproductive contexts, with both male and female gobies showing directed movement towards a calling male. The two-choice laboratory experiment was used (active versus quiet speaker) to analyze behavior of gobies in response to sound stimuli. When conspecific male spawning sounds were played, gobies moved in a direct path to the active speaker, suggesting true localization to sound. Of the animals that responded to conspecific sounds, 85% of the females and 66% of the males moved directly to the sound source. Auditory playback of natural and synthetic sounds showed differential behavioral specificity. Of gobies that responded, 89% were attracted to the speaker playing *Padogobius martensii* sounds, 87% to 100 Hz tone, 62% to white noise, and 56% to *Gobius niger* sounds. Swimming speed, as well as mean path angle to the speaker, will be presented during the presentation. Results suggest a strong localization of the round goby to a sound source, with some differential sound specificity.

8:30

3aABa3. Do laboratory rearing conditions affect auditory and mechanosensory development of zebrafish (*Danio rerio*)? Kirsten R. Poling, Eva Jaworski, Kristen R. Fantetti, and Dennis M. Higgs (Dept. of Biol., Univ. of Windsor, Windsor, ON, Canada N9B 3P4, kpoling@uwindsor.ca)

The effect of anthropogenic noise on the fish auditory system has become of increasing concern due to possible detrimental effects of intense sounds on auditory function and structures. This is especially problematic when raising fish in laboratory and aquaculture settings using filtration and aeration, which increase sound levels. To assess the effects of laboratory rearing conditions, one group of zebrafish (*Danio rerio*) embryos (“controls”) were placed into aerated aquaria in a normal laboratory rearing environment. A second set of embryos (“quiet”) were reared in aquaria with no aeration or filtration in a sound-resistant room. The intensity difference between the two sets of tanks was over 30 dB. Preliminary data show that there was no affect of differential rearing environments on saccular hair cell numbers or on hearing ability in fish up to 25 mm total length. However, rearing environment did affect neuromast number. “Quiet” fish had higher numbers of both cephalic and trunk superficial neuromasts, relative to controls. This difference was maintained up to 11 mm total length (the size at which canal formation begins). This suggests that acoustic environments normally found in the laboratory do not affect development of hearing in zebrafish, although laboratory acoustics may affect mechanosensory development.

8:45

3aABa4. The effect of sweep direction on avian auditory brainstem responses. Elizabeth Brittan-Powell, Amanda Lauer, Julia Callahan, Robert Dooling (Dept of Psych., Univ. of Maryland, College Park, MD 20742), Marjorie Leek (Army Medical Ctr., Washington, DC 20307), and Otto Gleich (Univ. of Regensburg, 93042 Regensburg, Germany)

In mammals, brief rising frequency sweeps result in increased amplitudes for both auditory brainstem response (ABR) and compound action potential (CAP) recordings (Dau, 2000; Shore and Nuttall, 1985). The rising sweep is thought to result in increased synchronous activity. Changing the direction of the sweep exaggerated the delay of processing along the basilar membrane and decreased synchrony of neural responses. Here we recorded ABRs from budgerigars, canaries, and zebra finches to a variety of stimulus parameters, including rising and falling sweeps with different sweep rates, determined by changing duration and frequency range. Both linear and nonlinear sweeps in frequency over time were tested. Results show that rising sweeps produce larger peak amplitudes, shorter latencies and changes in wave morphology such as a narrower wave 1 width than falling sweeps, suggesting greater synchrony of response to sweeps moving from low to high frequency. These data are consistent with mammalian results, but with a different time scale related to temporal characteristics of cochlear stimulation on the short basilar papilla in birds. [Work supported by NIH DC00198, DC001372, DC04664.]

3aABa5. Novel technique for rapid screening of tinnitus in rats.

Jeremy G. Turner, Thomas J. Brozoski, Jennifer L. Parrish, Carol A. Bauer, Larry F. Hughes, and Donald M. Caspary (Southern Illinois Univ. School of Medicine, P.O. Box 19620, Springfield, IL 62794-9620, jturner@siu.edu)

Measuring tinnitus in laboratory animals is difficult, involving weeks or months of operant training. Preliminary data suggest that rapid screening for tinnitus in rats can be accomplished using an unconditioned acoustic startle reflex. In control animals, a gap in an otherwise constant acoustic background inhibits a subsequent startle response to a sound impulse. If, however, the background signal is qualitatively similar to the animal's tinnitus, poorer detection of the gap and less inhibition of the startle might be expected. Fourteen animals with putative tinnitus at 10 kHz and 13 control animals were tested for gap detection using three different background signals: broadband noise, and filtered bandpass noise centered either at 16 kHz (15.5–16.5 kHz) or at their suspected tinnitus locus of 10 kHz (9.5–10.5 kHz). As predicted, animals with evidence of tinnitus exhibited significantly worse gap detection at 10 kHz, and were not significantly different than control animals at 16 kHz and broadband noise. These results suggest a new methodology for rapidly detecting tinnitus in individual animals. Equipment donated by Hamilton-Kinder Inc Behavioral Testing Systems in the memory of SIU graduate Dorothy Jean Kinder (Walker). [Work supported by NIH grants AG023910-01 (JT), DC4830 (TB & CB), and DC00151 (DC).]

9:15

3aABa6. Emotional learning modulates the rats sensitivity to transient changes in correlation between sounds.

Juan Huang, Liang Li, Zhigang Yang, Junli Ping (Dept. of Psych., Natl. Key Lab. on Machine Percept. Speech and Hearing Res. Ctr., Peking Univ., Beijing 100871, China), Xian Liu, Yixin Chen, and Xihong Wu (Peking Univ., Beijing 100871, China)

Humans are sensitive to small discrepancies between two sounds, and can detect a transient change in sound correlation. Here, prepulse inhibition of the startle reflex was used to examine rats sensitivity to transient changes in correlation between two correlated broadband noises, which were delivered by two spatially separated loudspeakers. The results show that either an uncorrelated noise fragment (UCNF, a drop of inter-sound correlation from 1.00 to 0 and then return to 1.00) or an anti-phase noise fragment (APNF, a drop of inter-sound correlation from 1.00 to -1.00 and then return to 1.00) could be detected by rats, since each of the changes in correlation could act as a prepulse stimulus to inhibit the startle reflex. The duration threshold for detecting the APNF was much lower than that for the UCNF. The detection of each of the changes in correlation was improved either by prolonged testing or by temporally pairing the UCNF or APNF with footshock. Thus similar to humans, rats also have the sensitivity to a sudden change in inter-sound correlation. Moreover, an increase of the sensitivity can be induced in rats either by repeated exposure to the change in correlation or by emotional learning.

3aABa7. The role of tragus on echolocating bat, *Eptesicus fuscus*.

Chen Chiu (Neurosci. and Cognit. Sci. Program, Univ. of Maryland, College Park, MD 20742) and Cynthia Moss (Univ. of Maryland, College Park, MD 20742)

Echolocating bats produce ultrasonic vocal signals and utilize the returning echoes to detect, localize and track prey, and also to avoid obstacles. The pinna and tragus, two major components of the bats external ears, play important roles in filtering returning echoes. The tragus is generally believed to play a role in vertical sound localization. The purpose of this study is to further examine how manipulation of the tragus affects a free-flying bat's prey capture and obstacle avoidance behavior. The first part of this study involved a prey capture experiment, and the bat was trained to catch the tethered mealworms in a large room. The second experiment involved obstacle avoidance, and the bat's task was to fly through the largest opening from a horizontal wire array without touching the wires. In both experiments, the bat performed the tasks under three different conditions: with intact tragus, tragus-deflection and recovery from tragus-deflection. Significantly lower performance was observed in both experiments when tragi were glued down. However, the bat adjusted quickly and returned to baseline performance a few days after the manipulation. The results suggest that tragus-deflection does have effects on both the prey capture and obstacle avoidance behavior. [Work supported by NSF.]

9:45

3aABa8. Minimum audible angles for aerial pure tones in a northern elephant seal (*Mirounga angustirostris*).

Marla M. Holt, Ronald J. Schusterman (UC Santa Cruz Long Marine Lab., 100 Shaffer Rd, Santa Cruz, CA 95060), Brandon L. Southall (UCSC Long Marine Lab. and NOAA Fisheries Acoust. Program, Silver Spring, MD 20910), and David Kastak (UC Santa Cruz Long Marine Lab., Santa Cruz, CA 95060)

Recent work has shown that several pinniped species localize aerial broadband signals as accurately as some terrestrial carnivores. Additionally, both harbor seals and California sea lions can better localize both the lower and higher frequencies of their hearing range compared to performance at intermediate frequencies. These results are congruent with the duplex theory of sound localization which states that low frequencies are localized by interaural time differences while high frequencies are localized by interaural intensity differences. Northern elephant seals are land breeding pinnipeds whose range of best hearing sensitivity is shifted toward lower frequencies compared to other pinnipeds tested thus far. In this study, we tested a female northern elephant seal in a hemi-anechoic chamber at six frequencies ranging between 0.8 and 16 kHz that were presented at levels approximately 25 dB above threshold. A left/right behavioral procedure was used to measure minimum audible angles (MAAs) at 75 percent correct discrimination. MAAs ranged from approximately three to fifteen degrees. Best performance occurred at the lower frequencies while worse performance occurred at the two highest test frequencies. Unlike sea lions and harbor seals, this subject showed a decreased ability to utilize interaural intensity differences above 4 kHz.

Session 3aABb

Animal Bioacoustics: Animal Acoustic Repertoires: Characterization and Classification

Volker Deecke, Chair

*Univ. of British Columbia, Marine Mammal Research Unit, 6248 BioSciences Rd., Vancouver, BC V6T 1Z4, Canada***Contributed Papers**

10:15

3aABb1. Variation in acoustic activity of North Atlantic right whales in three critical habitat areas in 2004. Susan E. Parks and Christopher W. Clark (Bioacoustics Res. Program, Cornell Univ., 159 Sapsucker Woods Rd., Ithaca, NY 14850)

The North Atlantic right whale is a critically endangered species. Five major habitat areas have been defined based on the presence of groups of right whales during different seasons of the year. Autonomous seafloor acoustic recorders, "pop-ups," have been used over the past five years to identify the presence/absence of right whales in Cape Cod Bay, Massachusetts by detection of right whale contact calls. These recordings have shown the effectiveness of using acoustics to monitor right whale behavior. In 2004, multi-unit arrays of pop-ups were deployed in three critical habitat areas; Cape Cod Bay (January–April), Great South Channel (May), and the Bay of Fundy (August). The presence/absence of right whale calls within range of the array, estimates of the minimum number of individuals calling, the types of calls recorded, and the diurnal patterns of sound production were measured to compare the acoustic repertoire use of right whales in these three habitats. These data were compared to visual survey data in the Bay of Fundy. These results can be used to compare diurnal and nocturnal behavioral patterns of the whales and provide information on behavioral differences between these three critical habitat areas.

10:30

3aABb2. Variation in acoustic behavior of delphinids in the Pacific Ocean based on school size and species composition. Shannon Rankin and Jay Barlow (SWFSC/NOAA, 8604 La Jolla Shores Dr., La Jolla, CA 92037, shannon.rankin@noaa.gov)

Variation in acoustic behavior based on school size and species composition was examined for surveys in the eastern tropical Pacific (2000), along the U.S. West Coast (2001), and in the U.S. EEZ surrounding Hawaii (2002). Sounds were monitored using a towed hydrophone array, and vocal schools were defined as those producing any combination of whistles, burst pulses, and/or echolocation clicks. Delphinid schools containing mixed species were consistently more vocal than single species schools. Vocal schools of *Stenella attenuata*, *S. longirostris*, *Delphinus delphis*, and *Lissodelphis borealis* were significantly larger than non-vocal schools. Vocal schools of *Tursiops truncatus* and *Grampus griseus* were somewhat larger than non-vocal schools, although this relationship was not significant. There was no relationship between group size and vocal activity for *S. coeruleoalba*, *Steno bredanensis*, and *Globicephala* spp. For species without a strong group size effect, all but *T. truncatus* were more vocal in the Hawaiian waters. The ability to use acoustic techniques in dolphin population estimation depends on their effectiveness in consistently detecting dolphin schools. This study suggests that small single-species schools of *S. attenuata*, *S. longirostris*, *D. delphis*, and *L. borealis* are more likely to be missed during acoustic monitoring in these regions.

10:45

3aABb3. Long term stability and individual distinctiveness in captive orca vocalizations. Michael Noonan and Malini Suchak (Canisius College, Buffalo, NY 14208, noonan@canisius.edu)

With focus on the question of signature calling in killer whales, recordings from five captive orcas (of Icelandic origin) held at Marineland of Canada were compared. For the present analysis, samples of three different call syllables were selected from recordings made five years apart and from instances in which the identity of the calling whale was unambiguous due to temporary isolation, concomitant bubbling, and/or head nodding. The Raven software package was used to ascertain the frequency range, frequency (max), duration, and timing of maximum and minimum power within each sample. For two of the three call syllables, statistically significant differences were found among the five whales for call length and for the timing of maximums and minimums ($p < 0.01 - 0.001$). This similarly proved true for nearly all pairwise comparisons between whales, including mother-offspring dyads. By contrast, for three of four whales for which we had sufficient samples, no significant differences were found on any measure between samples taken from the same whales five years apart. These findings therefore support the notion that the voices of individual orcas are distinct from one another in ways that are stable over the course of multiple years.

11:00

3aABb4. The acoustic repertoire of bottlenose dolphins (*Tursiops truncatus*) from the southern Gulf of Mexico. Carmen Bazua-Duran (Lab. Acustica, Facultad de Ciencias, UNAM, Cd. Universitaria, 04510 Mexico DF, Mexico)

Bottlenose dolphins live in a variety of habitats of the world's oceans using their acoustic repertoire to communicate and inspect their environment. This work investigates the acoustic repertoire of bottlenose dolphins that inhabit a coastal lagoon of the southern Gulf of Mexico, the Laguna de Terminos and how it may relate to the dolphins' general behavioral state and herd size, and to the general characteristics of the habitat, such as visibility, depth, and sea state. Preliminary results show that bottlenose dolphins produce by far more clicks than whistles in all behavioral states (feeding, resting, social, and travel) and herd sizes, which may correlate with the decreased visibility and shallow depth of the Laguna de Terminos. Additionally, silence was found during all behavioral states, but very seldom in herds of large size. These preliminary results suggest that bottlenose dolphins are choosing when and where to produce their phonations. Therefore, more detailed studies are needed to understand how these animals are using their acoustic sense to communicate and inspect their environment. [Work supported by CONACyT-Gobierno Edo. de Campeche and PAPIIT, UNAM.]

11:15

3aABb5. Automated categorization of bioacoustic signals: Avoiding perceptual pitfalls. Volker B. Deecke (Marine Mammal Res. Unit, Univ. of British Columbia, 6248 Biosciences Rd., Vancouver, BC, Canada V6T 1Z4) and Vincent M. Janik (Univ. of St. Andrews, St. Andrews, Fife KY16 8LB, Scotland, UK)

Dividing a species acoustic repertoire into biologically relevant categories presents a widespread problem in the study of animal sound communication. Such categorization is fundamental to any attempt to compare repertoires between contexts, individuals, populations or species. Automated procedures allow rapid, repeatable and objective categorization, however, in the past they have often performed poorly at detecting biologically meaningful sound classes. Arguably this is because automated methods have often failed to address the nonlinearities of animal sound perception and a method that incorporates dynamic time warping and an adaptive resonance theory (ART) neural network to avoid some of the problems is presented here. The method was tested on 104 randomly chosen whistles recorded from four captive bottlenose dolphins (*Tursiops truncatus*). This data set includes known biologically meaningful categories in the form of 42 distinctive stereotyped whistles produced when each individual was kept in isolation. The automated procedure correctly grouped all but two of the stereotyped whistles into their respective categories, and thus performed as good as human observers. However, compared to humans, the automated system provided finer categorization of the remaining non-stereotyped whistles. These results suggest that this methodology provides a repeatable and objective means of splitting bioacoustic signals into biologically meaningful categories.

11:30

3aABb6. Unsupervised classification of beluga whale vocalizations. Patrick J. Clemins and Michael T. Johnson (Marquette Univ., P.O. Box 1881, Milwaukee, WI 53201, patrick@clemins.name)

Unsupervised classification techniques are designed to discover natural groupings in a collection of data. For instance, unsupervised classification methods can be used to help define a repertoire for a species by grouping similar vocalizations together. The vocalizations grouped together are then defined as a single call type in the repertoire. There are a number of

difficulties in applying these algorithms to vocalizations including quantifying the vocalization into a set of features and accounting for temporal variations in the vocalizations. Using generalized perceptual linear prediction coefficients and a set of hidden Markov models to overcome these difficulties, beluga whale vocalizations are classified using *K*-means unsupervised classification. These classifications are then compared to labels assigned by human experts in order to determine the reliability of the classification system. The underlying goal of this research is to establish a generalized framework that can be used to analyze and classify the vocalizations of a number of species. [Work supported by the National Science Foundation under the Dr. Dolittle Project.]

11:45

3aABb7. Whale song analyses using bioinformatics sequence analysis approaches. Yian A. Chen, Jonas S. Almeida (Dept. of Biostatistics, Bioinformatics, and Epidemiology, 135 Cannon St., Medical Univ. of South Carolina, Charleston, SC 29425), and Lien-siang Chou (Inst. of Ecology & Evolutionary Biol., NTU, Taipei, Taiwan)

Animal songs are frequently analyzed using discrete hierarchical units, such as units, themes and songs. Because animal songs and bio-sequences may be understood as analogous, bioinformatics analysis tools DNA/protein sequence alignment and alignment-free methods are proposed to quantify the theme similarities of the songs of false killer whales recorded off northeast Taiwan. The eighteen themes with discrete units that were identified in an earlier study [Y. A. Chen, masters thesis, University of Charleston, 2001] were compared quantitatively using several distance metrics. These metrics included the scores calculated using the Smith-Waterman algorithm with the repeated procedure; the standardized Euclidean distance and the angle metrics based on word frequencies. The theme classifications based on different metrics were summarized and compared in dendrograms using cluster analyses. The results agree with earlier classifications derived by human observation qualitatively. These methods further quantify the similarities among themes. These methods could be applied to the analyses of other animal songs on a larger scale. For instance, these techniques could be used to investigate song evolution and cultural transmission quantifying the dissimilarities of humpback whale songs across different seasons, years, populations, and geographic regions. [Work supported by SC Sea Grant, and Ilan County Government, Taiwan.]

WEDNESDAY MORNING, 18 MAY 2005

PLAZA A, 9:00 TO 11:15 A.M.

Session 3aAO

Acoustical Oceanography, Underwater Acoustics and Animal Bioacoustics: Acoustics and Ocean Observatories II

James H. Miller, Chair

Dept. of Ocean Engineering, Univ. of Rhode Island, Narragansett Bay Campus, Narragansett, RI 02882

Invited Papers

9:00

3aAO1. Long-term variability in hydrothermal vent flow and temperature fluctuations by acoustic scintillation. Daniela Di Iorio (Dept. of Marine Sci., Univ. of Georgia, Athens, GA 30605)

An acoustical scintillation instrument is described that has been used to measure flow and temperature fluctuations at a hydrothermal vent plume. The vertical buoyancy driven flow together with the root-mean-square temperature fluctuations are obtained for vent structure 8F of the Main Endeavour vent field on the Endeavour segment of the Juan de Fuca Ridge. Long term (71 days) measurements are obtained and an estimate of the heat flux density which is of the order 0.06 MW m^{-2} is calculated. Measurements also show oscillations in the log-amplitude variance that result from either plume advection or increased entrainment of ambient fluids by the ambient tidal currents, thus demonstrating the need for a long time series measurement of simultaneous horizontal and vertical

flow. Future measurements in the Main Endeavour vent field, with possible integration into the NEPTUNE seafloor observatory, will be over extended periods of time so that comparisons of these processes can be made with independent seismic data collected by NOAA/PMEL SOSUS array.

9:20

3aAO2. Real-time observations and predictions of scour burial at the Martha's Vineyard Coastal Observatory. Michael D. Richardson (Marine Geosciences Div., Naval Res. Lab., Stennis Space Center, MS 39520, mrichardson@nrlssc.navy.mil), Peter Traykovski, Arthur C. Trembanis (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Carl T. Friedrichs (Virginia Inst. of Marine Sci., Gloucester Point, VA 23062)

Mine burial by scour was measured in real-time using cylindrical instrumented mines and rotary and two-axis pencil beam sonar. Two mines and the sonar were connected to a shore-based facility at the Martha's Vineyard Coastal Observatory. Data on mine movement (heading, pitch, and roll), scour pit geometry, percent burial, and environmental processes responsible for scour burial including significant wave height, period, and tidal height were analyzed daily and presented on a mine burial web site. Sonar images show scour pits developing around the mines in response to storm generated significant wave heights greater than 2.0 m. Mines pitch, change heading as they roll into scour pits. A comparison of the observed mine burial and real-time predictions using a modified version of the Whitehouse equations for wave-induced scour model exhibited good agreement suggesting mine burial by scour in similar sandy environments is predictable from bathymetry, sediment type, and measured or predicted surface wave conditions. The Coastal Observatory at Martha's Vineyard provided a unique opportunity for real-time data collection and analysis of both sediment transport events and seafloor-object interactions during storms. [Work supported by ONR.]

9:40

3aAO3. Autonomous and ship-cabled, bottom mounted sonar systems—Development, uses and issues associated with transitioning to ocean observatories. Kevin L. Williams, Russell D. Light, and Vernon W. Miller (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, williams@apl.washington.edu)

Three bottom mounted sonar systems will be described that were built over a span of fifteen years. The complexity of deployment and sophistication of the tasks performed increased with each system. The first system is an autonomous tower with rotating sonar designed to examine backscattering from an area within 50 m radius of the tower. The second is a ship-cabled system that includes a diver movable tower and separate buried array for examining both backscattering and acoustic penetration into sediments. The last is a ship-cabled rail/tower system designed to carry out forward scattering and synthetic aperture backscattering measurements. All three systems are designed to remain deployed for time periods up to a couple of months. After describing the systems, their deployment and some example results, recent efforts will be described that are aimed at transitioning these types of systems to cabled ocean observatories. The overall goal of the talk is to indicate both the level of complexity that can be envisioned for bottom mounted systems as well as the new issues that must be addressed in moving to cabled ocean observatories. [Work supported by ONR.]

10:00–10:15 Break

Contributed Papers

10:15

3aAO4. Development of interactive digital scientific acoustic systems for monitoring and assessment of physical and biological variables from ocean observatory platforms. Robert McClure, Tim Acker, James Dawson, and Assad E. K. Ebrahim (BioSonics, Inc., 4027 Leary Way NW, Seattle, WA 98107)

Scientific echosounders are in regular use for measurement of fish and plankton distribution and abundance, fish behavior, submersed aquatic vegetation, bottom sediment classification and water mass interface characteristics. Both fixed-aspect and mobile acoustic data are commonly collected from surface-deployed echosounders. Advancements in digital scientific echosounder technology and integration of other technologies has moved the capabilities of advanced echosounder systems into the realm of new applications and deployment possibilities. These advancements make possible the deployment of a new generation of integrated and interactive sensors which can provide high resolution, continuous, extended duration, recording, intelligent processing, and reliable monitoring and assessment of both biological and physical parameters of the ocean environment. Deployments from cabled observatories, moored and drifting buoys, manned and unmanned mobile platforms, and fully autonomous installations are all now reality. The deployment of these quality scientific instrument suites hold unlimited potential for research, resource management, and education. Existing and planned implementation of advanced digital scientific

echosounder applications on observatories as well as considerations in design, deployment, and integration with other observatory and oceanic acoustic activities are discussed.

10:30

3aAO5. Passive acoustic detection of fish sound production: The promise of ocean observatories. David Mann (USF College of Marine Sci., 140 7th Ave. S., St. Petersburg, FL 33701, dmann@marine.usf.edu)

Many fishes make species-specific sounds during courtship and spawning. These sounds can be recorded and quantified to study the temporal and spatial dynamics of spawning. Fish sounds typically range in frequency from 20–1500 Hz, and are usually inversely related to fish size. Time series of fish sound production recorded with autonomous recorders from Tampa Bay, Florida, and the mid-Atlantic bight, New Jersey illustrate how passive acoustics can be used to determine diel and seasonal timing of spawning. The great promise of ocean observatories is that they will not only provide unprecedented data on fish sound production, but that these data will be coupled with physical data on the same time and space scales. Ocean observatories will be powerful tools in studying the role of physical factors in driving variability in spawning and recruitment of fishes.

10:45

3aAO6. Basin Acoustic Seamount Scattering Experiment (BASSEX '04) III. Array performance. Kyle M. Becker (Penn. State Univ., Appl. Res. Lab., State College, PA 16804), Arthur B. Baggeroer (MIT, Cambridge, MA 02090), Kevin D. Heaney (Ocean Acoust. Services and Instrumentation Systems, Inc., Lexington, MA), Edward K. Scheer, and Keith von der Heydt (Woods Hole Oceanogr. Inst., Woods Hole, MA)

The ONR Five-Octave Research Array (FORA) was employed as the acoustic receiving system during the Basin Acoustic Seamount Scattering Experiment (BASSEX04). The array was primarily used in a towed configuration with receptions recorded for source/receiver separation distances ranging from 500–3200 km. Measurements made in proximity to the 75 Hz source moored near Kauai used a vertical array configuration. This paper presents array capabilities and performance during this first deep-water application of the FORA. Both acoustic and non-acoustic array performance during the BASSEX will be examined. In particular, array attitude and stability will be reviewed for different tow conditions, and for the array vertically deployed. Flow noise over the array will be examined for different tow speeds. Own-ship noise contamination will be presented for the array deployed vertically. Acoustic data will be presented as both individual channel measurements and beamformed output for a variety of source/receiver geometries and propagation conditions. Data quality will be assessed as it pertains to the stated objectives of a particular measure-

ment. Based on both acoustic and non-acoustic performance characteristics the advantages and/or deficiencies of the FORA will be discussed for its use in future low-frequency long-range acoustic propagation research. [Work supported by ONR Ocean Acoustics Program.]

11:00

3aAO7. Remote acoustic observations from a real-time observing system of the response of a shallow coastal embayment to hurricane forcing. Douglas J. Schillinger and Alex E. Hay (Dept. of Oceanogr., Dalhousie Univ., Halifax, NS, Canada B3H 4J1)

Results are presented from the acoustic remote sensing component of Dalhousie University's interdisciplinary Coastal Ocean Observatory in Lunenburg Bay, Nova Scotia, operational since June 2002 (www.cmep.ca/bay). Remote acoustic measurements of water column velocity and backscatter are made from bottom-mounted instruments hard-wired to surface buoys. The data are telemetered in near real-time to a shore station via wireless Ethernet, and to Dalhousie over the Internet. The fixed-point measurements from the bottom pods are augmented by a vessel-mounted profiling system, and by opportunistic deployments of internally-recording instruments. Highlights from the first three years of the observing system operation will be presented, including sediment suspension from the seabed and bubble penetration from the surface during Hurricane Juan.

WEDNESDAY MORNING, 18 MAY 2005

REGENCY A, 8:00 TO 9:45 A.M.

Session 3aBBa

Biomedical Ultrasound/Bioresponse to Vibration: Gene Therapy and Molecular Imaging

Michalakis A. Averkiou, Chair

Philips Ultrasound, 22100 Bothell-Everett Hwy., Bothell, WA 98041-3003

Chair's Introduction—8:00

Invited Papers

8:05

3aBBa1. Augmented and selective delivery of liquid perfluorocarbon nanoparticles to melanoma cells with ultrasound. Michael S. Hughes, Kathryn C. Crowder, Gregory M. Lanza, and Samuel A. Wickline (322 Yalem, 216 S. Kings Hwy., Washington Univ. Med. School, St. Louis, MO 63110, msh@cvu.wustl.edu)

Previous work has demonstrated the ability of liquid perfluorocarbon (PFC) nanoparticles to deliver therapeutic agents to cells selectively by binding to specific cellular epitopes, and confirmed the ability to simultaneously image these targeted nanoparticles with ultrasound. In this study, enhanced delivery of targeted PFC nanoparticles to cells expressing the integrin avb3 using clinical levels of ultrasound energy were studied. Nanoparticles complexed with ligands targeted to avb3 were incubated with cells (C32 melanoma) that expressed avb3 in culture. Control nanoparticles were produced that carried no ligand targeted to avb3. A custom specimen holder permitted simultaneous microscopic visualization of cell interactions during exposure to calibrated levels of ultrasound energy. After nanoparticle binding to cells and application of ultrasound, a roughly 2-fold increase in PFC content of the cells was observed. For control (nonbinding) nanoparticles, ultrasound exposure also increased PFC deposition, but the overall level was substantially less. Videodensitometric data show that nanoparticles were not destroyed by ultrasound exposure. Moreover, the alignment of nanoparticles relative to incident acoustic field demonstrate conclusively that acoustic radiation forces influence the nanoparticles and implicate these forces as participates in the enhanced delivery.

8:25

3aBBa2. The application of acoustic radiation force for molecular imaging and drug delivery. Paul A. Dayton and Katherine W. Ferrara (UC Davis Biomed. Eng., 1 Shields Ave., Davis, CA 95616)

Acoustic radiation force is exerted on objects in an acoustic field. By optimizing acoustic parameters to maximize this force, ultrasound contrast agents can be manipulated and concentrated with an acoustic field. This technique can be used to enhance molecular imaging and drug delivery. Ultrasonic molecular imaging employs contrast agents such as microbubbles, nanoparticles, or liposomes coated with ligands specific for receptors expressed on cells at sites of angiogenesis, inflammation, or thrombus. Concentration of these highly echogenic contrast agents at a target site enhances the ultrasound signal received from that site, promoting

ultrasonic detection and analysis of disease states. It is demonstrated that acoustic radiation force can be used to displace targeted contrast agents to a vessel wall, greatly increasing the number of agents binding to available surface receptors. Radiation force can also be used for microparticle-carrier based drug delivery. Acoustically active drug-carrier vehicles can be concentrated with radiation force, and then disrupted with high-intensity ultrasound bursts, resulting in local delivery of the therapeutic agent. This talk will discuss simulations and experimental results demonstrating these applications.

Contributed Papers

8:45

3aBBa3. A new targeted drug delivery method using ultrasound and acoustically active lipospheres. Michaelann Shortencarier, Susannah Bloch, Paul Dayton, Kathy Ferrara (Dept. of Biomed. Eng., UC Davis, 451 East Health Sci. Dr., Davis, CA 95616), Terry Matsunaga, Rachel Labell, and Patricia Schumann (ImaRx Therapeutics, Inc., Tucson, AZ 85719)

The goal is to create a strategy for localized drug delivery using engineered delivery vehicles and ultrasound energy. These drug delivery vehicles, referred to as acoustically active lipospheres (AALs), consist of small gas bubbles surrounded by thick oil shells (where drugs can be carried) and are enclosed by an outermost lipid layer. Ultrasound radiation force can be used to displace these vehicles near the blood vessel wall, after which a higher intensity pulse can fragment the vehicle and transfer its contents to the endothelium. Blood velocity ranges from 1 to 10 mm/s in the microvasculature. Therefore, agents in capillaries insonified by a transducer with a 1 mm focal beam width for 0.1 to 1 seconds can be displaced of 1 to 2 mm. Delivery vehicles containing various fluorescent dyes within the oil layer are exposed to ultrasound in cell chambers or flowing vessels. Fluorescence intensity increases more than 10-fold with application of radiation force and fragmentation, both over no exposure and over fragmentation pulses alone (both $P < 0.001$). This implies radiation force is necessary to bring AALs into proximity of the cell monolayer before their destruction in order for drug delivery to occur.

9:00

3aBBa4. Delivery of anticancer drugs and antibodies into cells using ultrasound. Junru Wu, Jason Pepe (Dept. of Phys., Univ. of Vermont, Burlington, VT 05405), and Mercedes Rincon (Univ. of Vermont, Burlington, VT 05405)

It has been shown experimentally in cell suspensions that pulsed ultrasound (2.0 MHz) could be used to deliver an anti-cancer drug (Adriamycin hydrochloride) into Jurkat lymphocytes and antibodies (goat anti rabbit IgG and anti mouse IgD) into human peripheral blood mononuclear (PBMC) cells and Jurkat lymphocytes assisted by encapsulated microbubbles (Optison). When Adriamycin hydrochloride (ADR) was delivered, the delivery efficiency reached 4.80% and control baseline (no ultrasound and no ADR) was 0.17%. When anti-rabbit IgD was delivered, the efficiencies were 34.90% (control baseline was 1.33%) and 32.50% (control baseline was 1.66%) respectively for Jurkat cells and PBMC. When goat anti rabbit IgG was delivered, the efficiencies were 78.60% (control baseline was 1.60%) and 57.50% (control baseline was 11.30%) respectively for Jurkat cells and PBMC.

9:15

3aBBa5. 1 kHz vibration increases proteoglycan production in ATDC5 chondrocytes. Heather M. Argadine, Randall R. Kinnick, James F. Greenleaf (Dept. of Physio. and Biomed. Eng., Mayo Clinic College of Medicine, 200 First St. SW, Rochester, MN 55905, argadine.heather@mayo.edu), and Mark E. Bolander (Mayo Clinic College of Medicine, Rochester, MN 55905)

In vitro studies have shown that treatment with 1.5 MHz ultrasound signal (160 mW/cm²) at a 200 μ s tone burst repeating at 1 kHz increases proteoglycan synthesis in chondrocytes [J. Parvisi *et al.*, J. Orthop. Res. **17**, 488–494 (1999)]. It was hypothesized that a continuous 1 kHz signal would be similar to the pulsed 1.5 MHz signal in stimulating chondrocytes to produce proteoglycan, which may cause accelerated fracture healing. *In vitro* experiments were performed with ATDC5 cells, a chondrogenic clonal cell line, plated in 6-well plates for 3 to 7 days before receiving ultrasound treatments. Cells were treated with either 1.5 MHz pulsed signal or 1 kHz signal for 20 minutes per day for 9 to 11 days. The signals were calibrated so that the bottom of the 6-well plate moved 10 nm for each condition. After the final treatment, cell layers were stained with Alcian blue, which stains cartilage nodules providing a measure of chondrogenesis. Both 1.5 MHz and 1 kHz led to a highly significant increase in chondrogenesis compared to control. Quantitative image analysis of stained wells showed that treatments with either signal increased number of nodules 2.3-fold ($p < 0.02$) and total area of nodules 3-fold ($p < 0.02$) compared to controls.

9:30

3aBBa6. A theoretical study of the feasibility of acoustical tweezer: Ray acoustics approach. Jungwoo Lee and Kirk Shung (Dept. of Biomed. Eng., Univ. of Southern California, DRB132, Los Angeles, CA 90089, jungwool@usc.edu)

Optical tweezer has been found to have many biomedical applications in trapping macromolecules and cells. For the trapping mechanism, there has to be a sharp spatial change in axial optical intensity and the particle size must be much greater than the wavelength. Similar phenomenon may exist in acoustics. This work was undertaken to demonstrate theoretically that it is possible to acoustically trap particles near the focal point if certain conditions are met. Acoustic force exerted on fat tissue in ultrasonic fields is analyzed in ray acoustics regime where the wavelength of acoustic beam is much smaller than the size of the particle. In this paper, the analysis is therefore based on the field pattern produced by a strongly focused 100 MHz ultrasonic transducer with Gaussian intensity distribution. The magnitude of force and Fresnel coefficients at various positions are calculated. According to the simulation results, acoustical tweezer works particularly when the beam width at focus is one wavelength and the tolerance of acoustic impedance mismatch between two media lies within 6.7%. [Work supported by NIH Grant P41-EB2182.]

Session 3aBBb

Biomedical Ultrasound/Bioresponse to Vibration: Contrast Agents and Cavitation

Paul A. Dayton, Chair

Univ. of California, Biomedical Engineering, 1 Shields Ave., Davis, CA 95616

Contributed Papers

10:15

3aBBb1. Cavitation effects of therapeutic ultrasound. Kenneth Bader and Karen Gipson (Grand Valley State Univ., Allendale, MI 49401-9403)

In order to model possible effects of cavitation on bone structure, the formation of cavitation along an acoustically reflective PVC surface was studied. The bubbles were generated by a commercial therapeutic transducer (1 MHz), for both water and agar gel mediums. Bubble formation and activity were monitored using light scattering techniques: a HeNe laser was directed through the medium, and the transmitted light was detected using an optical power meter. The data show a decline in cavitation activity further from the surface and offer suggestive evidence of a boundary layer. Temporal saturation is also evident in both mediums, though an anomalous increase in transmitted light for the agar medium suggests the possibility of pressure variations due to microstreaming. Analysis of the microstreaming pattern was subsequently performed at lower frequencies in order to allow visualization with a standard laboratory microscope.

10:30

3aBBb2. Interacting bubble dynamics in Hamiltonian mechanics. Yurii A. Ilinskii, Mark F. Hamilton, and Evgenia A. Zabolotskaya (ARL, Univ. of Texas, 10000 Burnet Rd., Austin, TX 78758-4423)

A Hamiltonian formulation of the equations of motion is presented for the oscillation and 3D translation of N interacting spherical bubbles. Existing alternative formulations contain second derivatives of both the bubble radius and the three coordinates of the position vector. For two or three bubbles the systems may be reduced analytically to explicit first-order equations that can be solved with standard numerical algorithms. For large N , the equations cannot realistically be recast in explicit first-order form, and the leading derivatives must be evaluated by matrix inversion at each time step. The inversions introduce substantial errors for large N . We avoid the inversions by using Hamiltonian mechanics to derive a system of explicit first-order equations for all N , and which includes the influence of an applied sound field. For large N , the approach based on Hamiltonian mechanics results in both increased numerical accuracy and reduced computation time. The Hamiltonian system is validated by comparison with solutions obtained by alternate means. Examples are presented demonstrating various effects of bubble interaction. Although developed for modeling cavitation in shock-wave lithotripsy, the model is also applicable to bubble clouds in the ocean. [Work supported by NIH Grant EB004047 and ARL IR&D funds.]

10:45

3aBBb3. Modeling of initial bubble growth rates during high-intensity focused ultrasound. Wayne Kreider, Michael R. Bailey, and Lawrence A. Crum (Appl. Phys. Lab, Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, wkreider@u.washington.edu)

In therapeutic applications of biomedical ultrasound, it is important to understand the behavior of cavitation bubbles. For applications that use high-intensity focused ultrasound (HIFU), both large negative acoustic pressures and heating can independently lead to bubble formation. Although neglected previously, heating during HIFU is expected to affect the growth and dissolution of bubbles by both raising the vapor pressure and

promoting outgassing from gas-saturated tissues. Herein, the dynamics of a single, spherical bubble in water have been modeled using the Gilmore equation closed with an energy balance on bubble contents for calculation of pressures inside the bubble. Moreover, heat and mass transfer at the bubble wall are incorporated using the Eller-Flynn zeroth-order approximation for gas diffusion, an estimation of non-equilibrium phase change based on the kinetic theory of gases, and assumed shapes for the spatial temperature distribution in the surrounding liquid [Yasui, J. Phys. Soc. Jpn. **65**, 2830–2840 (1996)]. This model allows explicit coupling of the ambient heating during HIFU to the thermodynamic state of an oscillating bubble and is currently being used to explore the growth rates of initially small, undetectable bubbles exposed to various HIFU treatment protocols. [Work supported by NIH T32-EB-001650, NIH DK43881, and NSBRI SMS00203.]

11:00

3aBBb4. Microbubble dynamics in contrast agent shell rupture. Azzdine Y. Ammi, S. Lori Bridal (CNRS UMR 7623, Université Pierre et Marie Curie, 15 rue de l'école de médecine, 75006 Paris, France, ammi@lip.bhdc.jussieu.fr), Robin O. Cleveland (Boston Univ., Boston, MA 02215), Jonathan Mamou, Grace I. Wang, and William D. O'Brien, Jr. (Univ. of Illinois, Urbana, IL 61801)

A passive cavitation detector (PCD) was used to study the features of insonified contrast agent microbubble rupture. When Optison microbubbles ruptured, the PCD signal received from single microbubbles showed the presence of post-excitation echoes identified as inertial cavitation and rebounds. Using these signals, the collapse threshold was evaluated as a function of insonation frequency, pulse duration, and pressure amplitude. For each excitation signal, 128 received time traces were recorded (200 MHz sampling, 12-bit), and frequency-domain analysis of the received signals was performed to characterize microbubble collapse. The peak rarefactional pressure threshold for detection of these emissions increased with frequency (e.g., 0.53, 0.87, and 0.99 MPa for 0.9, 2.8, and 4.6 MHz, respectively; 5-cycle pulse duration) and decreased with pulse duration. The modified Herring equation was used to model the bubble response at experimentally determined threshold conditions. These predicted microbubble dynamics were used to determine microbubble rupture indices based on: radial strain, peak velocity, and peak acceleration. This work provides a clearer understanding of the role of contrast agent dynamics in bubble destruction. [Work supported by the cooperative project for biomedical engineering between the University of Illinois at Urbana-Champaign, USA and the Centre National de la Recherche Scientifique, France.]

11:15

3aBBb5. The effect of size on the acoustic response of polymer-shelled ultrasound microbubble contrast agents. Susannah H. Bloch, Katherine W. Ferrara (Dept. of Biomed. Eng., Univ. of California Davis, Davis, CA 95616, kwferrara@ucdavis.edu), and Erik R. Wisner (Univ. of California Davis, Davis, CA 95616)

Hard-shelled microbubble contrast agents differ from lipid-shelled agents in that the extra stability provided by their shell prevents dissolution of small gas bubbles: consequently, stable agents can be manufactured with diameters below 1 μm . Near-micron and submicron-sized contrast

agents may enable new biomedical applications because they can traverse narrow barriers such as the lymphatic endothelium and the permeable vascular endothelium found in actively growing solid tumors. However, several studies have suggested fundamental differences in the mechanisms of acoustic response of “soft” lipid-shelled ultrasound microbubble contrast agents and “hard” albumin- or polymer-shelled agents. In particular, hard-shelled contrast agents exhibit strong echoes only after the shell is damaged and the enclosed gas is allowed to oscillate freely. Here three polymer-shelled agents with mean diameters of 0.74 μm , 0.91 μm , and 1.33 μm (POINT Biomedical Corp.) are investigated. Significant differences among the agents exist in the probability of response to acoustic interrogation, but not the amplitude or frequency of their acoustic response. These results are explained by a hypothesis of a pressure-dependent threshold diameter above which single hard-shelled agents respond to acoustic interrogation. The implications of this hypothesis for the design of imaging methods and clinical protocols using these agents will be presented.

11:30

3aBBb6. Contrast-enhanced detection of arterial bleeding. Wenbo Luo, Vesna Zderic, Frank Starr, Marla Paun (Appl. Phys. Lab., Univ. of Washington, Box 355640, 1013 NE 40th St., Seattle, WA 98105, wenbo@u.washington.edu), and Shahram Vaezy (Univ. of Washington, Seattle, WA 98105)

The purpose of this study was to determine the feasibility and usefulness of contrast agent to detect internal bleeding. In a model of acute femoral artery injury, bleeding sites were produced percutaneously in 9 rabbits, using an 18 G needle guided by Doppler and B-mode ultrasound. Imaging was performed pre- and post-injury and after i.v. injection of contrast agent (Optison) using B-mode, Color Doppler (CD), and Pulse Inversion. Gross and histological examination (light microscopy) of the injured blood vessel and the surrounding tissue were performed. Experiments with similar procedure were achieved in a pulsatile phantom with carotid arteries as targeted blood vessels in vitro. Bleeding sites could not be located in non-contrast B-mode sonography. After administration of contrast agent, injuries could be detected by the increased echogenicity of

the extravasated blood flow. The resulted hematoma appeared as inhomogeneous hyper-echo region. Color mosaic patterns were observed in the surrounding tissue with CD. Cluster of cavities were found within hematoma upon microscopy, which may be due to aggregated microbubbles. Bleeding sites were recognized in all contrast studies by CD; compared with only 11% in non-contrast control. Contrast-enhanced sonography is a promising tool in detection and evaluation of internal bleeding.

11:45

3aBBb7. Ultrasound contrast agents for bleeding detection and acoustic hemostasis. Vesna Zderic, Wenbo Luo, Andrew Brayman, Lawrence Crum (Appl. Phys. Lab. and Bioengineering, Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, vesna@u.washington.edu), and Shahram Vaezy (Univ. of Washington, Seattle, WA 98105)

Objective: To investigate the application of ultrasound contrast agents (UCA) in improving both therapeutic and diagnostic aspects of ultrasound-guided High Intensity Focused Ultrasound (HIFU) therapy. Methods: Incisions (3 cm long, 0.5 cm deep) were made in rabbit livers (in anterior surface for HIFU treatment, or posterior surface for bleeding detection). UCA Optison (~0.1 ml/kg) was injected into mesenteric vein or ear vein. A HIFU applicator (5.5 MHz, 6400 W/cm²) was scanned manually over the incision until hemostasis was achieved. Occult bleeding was monitored with Doppler ultrasound. Results: The presence of Optison produced 37% reduction in hemostasis times normalized to initial bleeding rates. Gross and histological observations showed similar appearance of HIFU lesions produced in the presence of Optison and control HIFU lesions. The temperature reached 100°C in both HIFU only and HIFU +UCA treatments. Tension strength of hemostatic liver incisions was 0.9±0.5 N. Almost no bleeding could be detected before Optison injection. First appearance of contrast enhancement localized at the bleeding site was 15 s after Optison injection, and lasted for ~50 s. Conclusion: The presence of UCA during HIFU treatment of liver incisions resulted in shortening of HIFU application times and better visualization of bleeding sites.

WEDNESDAY MORNING, 18 MAY 2005

BALMORAL, 8:00 A.M. TO 12:00 NOON

Session 3aMU

Musical Acoustics: General Topics in Musical Acoustics (Lecture/Poster Session)

Gary P. Scavone, Cochair

McGill Univ., Music Technology, Faculty of Music, Montreal, QC H3A, 1E3, Canada

R. Dean Ayers, Cochair

Southern Oregon Univ., 1250 Siskiyou Blvd., Ashland, OR 97520

Contributed Papers

8:00

3aMU1. Phase preservation in musical signals. V. Vijayakumar and C. Eswaran (Multimedia Univ., Jalan Air Keroh Lama, 75450 Melaka, Malaysia)

The intensity variations of the harmonics of musical instruments in the frequency domain can be interpreted to store phase information. The motivation for this arises by considering the similarity of the timbre of in-

struments to that obtained through a diffraction model proposed here. The intensity modifications of an input spectrum of discrete frequencies of unit intensity into the known spectra of different musical instruments have been found to match that of known instruments. According to diffraction theory, the modifying envelope encodes phase information. By considering the similarity, it is proposed that musical instrument timbre store phase information. It is suggested that timbre itself could have diffraction origins. Specific examples of musical instruments are considered to illustrate this interpretation.

3aMU2. Practical system for recording spatially lifelike 5.1 surround sound and 3D fully periphonic reproduction. Robert E. (Robin) Miller III (FilmmakerTechnol., 606 W. Broad St., Bethlehem, PA 18018)

In acoustic spaces that are played as extensions of musical instruments, tonality is a major contributor to the experience of reality. Tonality is described as a process of integration in our consciousness over the reverberation time of the room of many sonic arrivals in three dimensions, each directionally coded in a learned response by the listeners unique head-related transfer function (HRTF). Preserving this complex 3D directionality is key to lifelike reproduction of a recording. Conventional techniques such as stereo or 5.1-channel surround sound position the listener at the apex of a triangle or the center of a circle, not the center of the sphere of lifelike hearing. A periphonic reproduction system for music and movie entertainment, Virtual Reality, and Training Simulation termed PerAmbio 3D/2D (Pat. pending) is described in theory and subjective tests that capture the 3D sound field with a microphone array and transform the periphonic signals into ordinary 6-channel media for either decoderless 2D replay on 5.1 systems, or lossless 3D replay with decoder and five additional speakers. PerAmbio 3D/2D is described as a practical approach to preserving the spatial perception of reality, where the listening room and speakers disappear, leaving the acoustical impression of the original venue.

8:30

3aMU3. The glissando illusion: A spatial illusory contour in hearing. Diana Deutsch, Kamil Hamaoui (Dept. of Psych., Univ. of California, San Diego, La Jolla, CA 92093, ddeutsch@ucsd.edu), and Trevor Henthorn (Univ. of California, San Diego, La Jolla, CA 92093)

In the glissando illusion (originally demonstrated by Deutsch, 1995) a synthesized oboe tone of constant pitch is played together with a sine wave whose pitch repeatedly glides up and down. These are presented through two loudspeakers, placed to the listeners left and right, and are repeatedly switched between loudspeakers, such that when the oboe tone emanates from one, a segment of the sine wave emanates from the other. The oboe tone is perceived correctly as switching between loudspeakers; however the segments of the sine wave appear to be joined together seamlessly, such that a single, continuous tone is heard, which appears to emanate from a source that moves slowly around in space in accordance with its pitch motion. In the present experiment, righthanders ($n=22$) tended strongly to hear the glissando move between left and right, and also between low and high in space, as its pitch moved between low and high. It was therefore frequently heard as tracing an elliptical path aligned diagonally between a position low and to the left when its pitch was lowest, and high and to the right when its pitch was highest. Nonrighthanders ($n=42$) perceived the illusion in statistically different ways.

8:45

3aMU4. Effect of critical-band smoothing of musical instrument spectral data. James W. Beauchamp (School of Music and Dept. of Elec. and Computer Eng., Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801, jwbeauch@uiuc.edu) and Andrew B. Horner (Hong Kong Univ. of Sci. and Technol., Clear Water Bay, Kowloon, Hong Kong)

It has been found that second-order harmonic smoothing of musical instrument spectral data can have a significant effect on timbral perception, depending on the instrument tested [McAdams *et al.*, *J. Acoust. Soc. Am.* **102**, 882–897 (1999)]. With critical-band smoothing, the lower harmonics, since they are in different critical bands, retain their individual amplitudes and temporal envelopes. Thus, it is hypothesized that critical-band smoothing has a lesser perceptual effect on most instrument tones than harmonic smoothing. On the other hand, upper critical bands consist of groups of harmonics. It is hypothesized that it is difficult to hear out individual harmonics within critical bands. Thus, for each band the independent harmonic temporal envelopes can be replaced by a composite rms-amplitude envelope. Spectra within bands can be replaced by time-averaged spectra. Alternatively, time-dependent amplitude versus Bark-

frequency spectral envelopes can be smoothed for each individual analysis frame. Further, amplitudes can be averaged in dB or linear units. Results for various processing combinations and various musical instrument sounds will be given and demonstrated.

9:00

3aMU5. An improvisation on the Middle-Eastern mijwiz; auditory roughness profiles and tension/release patterns. Pantelis N. Vassilakis (DePaul Univ., School of Music, 804 W. Belden Ave., Chicago, IL 60614-3250, pantelis@acousticslab.com)

Within western musical tradition, auditory roughness constitutes one of the principle perceptual correlates of dissonance. A previous study [P. N. Vassilakis, *Selected Reports in Ethnomusicology*, Vol. 12 (in press)] addressed the musical significance of auditory roughness within western as well as three non-western musical traditions. It introduced a new roughness estimation model that predicts successfully roughness ratings of synthesized complex diatonic dyads. For the present study, an application was created that uses the aforementioned model to estimate the roughness of natural signals with arbitrary spectra, as well as roughness profiles of short musical pieces. The application was used to examine if and how the pattern of tension and release within a stylized improvisation on the Middle Eastern mijwiz (double clarinet) relates to roughness changes. Tension/release patterns were indicated by the Lebanese improviser and expert mijwiz player (Dr. A. J. Racy) and were also obtained by both musicians and non-musicians (American-raised) in a perceptual experiment. The results suggest that auditory roughness is a good predictor of the tension/release pattern indicated by the improviser. The patterns obtained by the subjects, although not very consistent, appear to be different overall, suggesting that musical tension and release may be culture-specific concepts.

9:15

3aMU6. Visual feedback for retuning to just intonation intervals. R. Dean Ayers (Phys. Dept., Southern Oregon Univ., 1250 Siskiyou Blvd., Ashland, OR 97520), Peter R. Nordquist, and Justin S. Corn (Southern Oregon Univ., Ashland, OR 97520)

Musicians become used to equal temperament pitch intervals due to their widespread use in tuning pianos and other fixed-pitch instruments. For unaccompanied singing and some other performance situations, a more harmonious blending of sounds can be achieved by shifting to just intonation intervals. Lissajous figures provide immediate and striking visual feedback that emphasizes the frequency ratios and pitch intervals found among the first few members of a single harmonic series. Spirograph patterns (hypotrochoids) are also especially simple for ratios of small whole numbers, and their use for providing feedback to singers has been suggested previously [G. W. Barton, Jr., *Am. J. Phys.* **44**(6), 593–594 (1976)]. A hybrid mixture of these methods for comparing two frequencies generates what appears to be a three dimensional Lissajous figure—a cylindrical wire mesh that rotates about its tilted vertical axis, with zero tilt yielding the familiar Lissajous figure. Sine wave inputs work best, but the sounds of flute, recorder, whistling, and a sung “oo” are good enough approximations to work well. This initial study compares the three modes of presentation in terms of the ease with which a singer can obtain a desired pattern and recognize its shape.

9:30

3aMU7. Directivity measurements of the violin during performance. Jean-François Hardy and Jean-Gabriel Migneron (Laboratoire d'acoustique de l'Université Laval, 1 Côte de la Fabrique, Québec, PQ, Canada G1K 7P4, jean-francois.hardy.1@ulaval.ca)

Many studies have shown that directivity of musical instruments is significant for the perceived sound within the hall and for performer's ease of hearing each other. The directivity patterns explain in part the differences between many common or special orchestral stage plots. Measure-

ments of the violin's directivity have been performed using an acoustical intensimetry procedure. The precise directivity characteristics of isolated tones have been compared with the mean directivity resulting from performance of four extracts chosen among the orchestral repertoire. Results, which were measured in both horizontal and vertical planes, show interesting differences between the average directivity of each test. They lead to links with few important distinctions in the traditional violin's positions inside the orchestra.

9:45

3aMU8. Application of cochlear analysis techniques to percussion in electro-acoustic music. John Mills III (ECE Dept, The Univ. of Texas, 1 Univ. Station 0803, Austin, TX 78712, nodog@mail.utexas.edu)

This presentation is focused upon the audio analysis involved in the creation of a "pseudo-score" for electro-acoustic music. This research was undertaken because the lack of an objective visual representation prevents music theorists from being able to easily discuss electro-acoustic music in written form. The "holy grail" of automatic computer music analysis has most often been the translation of a continuous pressure variation into traditional Western musical notation. Since electro-acoustic music is rarely able to be transcribed into this type of notation, the automatic analysis of electro-acoustic music confounds many previous algorithms. A top-down approach is suggested in order to extract acoustic and musical information from recordings of electro-acoustic music. Because humans are the best known system for translating a continuous pressure variation into useful acoustic information, a cochlear model, Patterson's Auditory Image Model (AIM), is used as a front end. The audio property of self-similarity of percussion has been chosen for initial study. This self-similarity property requires first that the audio property of percussion is quantified. The nuances of using AIM to determine percussion will be discussed. Models of self-similarity of percussion for varying time-scales will then be discussed.

10:45–10:50 Break

All posters will be on display and authors will be at their posters from 10:50 a.m. to 12:00 noon.

3aMU11. Microrhythmic characteristics of musical instrument initial transients. Rolf Bader (Inst. of Musicology, Univ. of Hamburg, Neue Rabenstr. 13, 20354 Hamburg, Germany)

A microrhythmic analysis of initial transients of the violin, the guitar, the saxophone, the clarinet, and percussion instruments has been performed. They show characteristic patterns between instrument families, concerning the derivation of microrhythmic intervals within the tonal frequency range compared to psychoacoustic discrimination tasks by listeners. Microrhythm here is defined as the temporal distance between two amplitude peaks. As initial transients are a crucial part of the sound in terms of identification of instruments and the initial is just in the range of the second integration time of the ear, where frequency components can just be identified, these microrhythmic patterns could be a possibility for listeners to identify the instrument families. This is because microrhythmic intervals are instantaneously present and do not need an integration time span.

10:00–10:15 Break

10:15

3aMU9. Fractal dimension analysis of complexity in Ligeti piano pieces. Rolf Bader (Inst. of Musicology, Univ. of Hamburg, Neue Rabenstr. 13, 20354 Hamburg, Germany)

Fractal correlation dimensional analysis has been performed with whole solo piano pieces by Gyrgy Ligeti at every 50ms interval of the pieces. The resulting curves of development of complexity represented by the fractal dimension showed up a very reasonable correlation with the perceptual density of events during these pieces. The seventh piece of Ligeti's "Musica ricercata" was used as a test case. Here, each new part of the piece was followed by an increase of the fractal dimension because of the increase of information at the part changes. The second piece "Galamb borong," number seven of the piano Etudes was used, because Ligeti wrote these Etudes after studying fractal geometry. Although the piece is not fractal in the strict mathematical sense, the overall structure of the psychoacoustic event-density as well as the detailed event development is represented by the fractal dimension plot.

10:30

3aMU10. Attack transients on lip reed instruments. Murray Campbell, Seona Bromage (School of Phys., Univ. of Edinburgh, Edinburgh EH9 3JZ, UK, d.m.campbell@ed.ac.uk), and John Chick (Univ. of Edinburgh, Edinburgh EH9 3JL, UK)

The generation of sound in a lip reed (brass) instrument relies on a non-linear coupling between two vibrating systems: the lips of the player, and the air column contained by the walls of the instrument. The nature of this coupling has been studied by filming the motion of the lips of a number of players using a specially constructed transparent mouthpiece, while simultaneously recording the pressure in the mouthpiece and at the bell of the instrument. Attention has been focused on the attack transient, and the way in which its duration and evolution depend on the extent to which the air column modes are harmonically related. For this reason the instrument chosen was the serpent; on this instrument certain fingerings give air column modes whose relationship is close to harmonic, while other fingerings give strikingly inharmonic air column modes. The relationship between measurements of transient character and judgements of the playability or responsiveness of the instrument will be discussed.

3aMU12. Signal-to-background-ratio preferences of normal-hearing listeners as a function of music. Jillian G. Barrett (Dept. of Communicative Sci. and Dis., California State Univ. Hayward, 25800 Carlos Bee Blvd., Hayward, CA 94542)

The primary purpose of speech is to convey a message. Many factors affect the listener's overall reception, several of which have little to do with the linguistic content itself, but rather with the delivery (e.g., prosody, intonation patterns, pragmatics, paralinguistic cues). Music, however, may convey a message either with or without linguistic content. In instances in which music has lyrics, one cannot assume verbal content will take precedence over sonic properties. Lyric emphasis over other aspects of music cannot be assumed. Singing introduces distortion of the vowel-consonant temporal ratio of speech, emphasizing vowels and de-emphasizing consonants. The phonemic production alterations of singing make it difficult for even those with normal hearing to understand the singer. This investigation was designed to identify singer-to-background-ratio (SBR) prefer-

ences for normal hearing adult listeners (as opposed to SBR levels maximizing speech discrimination ability). Stimuli were derived from three different original songs, each produced in two different genres and sung by six different singers. Singer and genre were the two primary contributors to significant differences in SBR preferences, though results clearly indicate genre, style and singer interact in different combinations for each song, each singer, and for each subject in an unpredictable manner.

3aMU13. Normal-hearing listener preferences of music as a function of signal-to-noise-ratio. Jillian G. Barrett (Dept. of Communicative Sci. and Dis., California State Univ. Hayward, 25800 Carlos Bee Blvd., Hayward, CA 94542)

Optimal signal-to-noise ratios (SNR) for speech discrimination are well-known, well-documented phenomena. Discrimination preferences and functions have been studied for both normal-hearing and hard-of-hearing populations, and information from these studies has provided clearer indices on additional factors affecting speech discrimination ability and SNR preferences. This knowledge lends itself to improvements in hearing aids and amplification devices, telephones, television and radio transmissions, and a wide arena of recorded media such as movies and music. This investigation was designed to identify the preferred signal-to-background ratio (SBR) of normal-hearing listeners in a musical setting. The signal was the singer's voice, and music was considered the background. Subjects listened to an unfamiliar ballad with a female singer, and rated seven different SBR treatments. When listening to melodic motifs with linguistic content, results indicated subjects preferred SBRs similar to those in conventional speech discrimination applications. However, unlike traditional speech discrimination studies, subjects did not prefer increased levels of SBR. Additionally, subjects had a much larger acceptable range of SBR in melodic motifs where the singer's voice was not intended to communicate via linguistic means, but by the pseudo-paralinguistic means of vocal timbre and harmonic arrangements. Results indicate further studies investigating perception of singing are warranted.

3aMU14. Soprano and source: A laryngographic analysis. Laura Anne Bateman (Canadian College of Performing Arts, 1701 Elgin Rd., Victoria, BC, Canada)

Popular music in the 21st century uses a particular singing quality for female voice that is quite different from the trained classical singing quality. Classical quality has been the subject of a vast body of research, whereas research that deals with non-classical qualities is limited. In order to learn more about these issues, the author chose to do research on singing qualities using a variety of standard voice quality tests. This paper looks at voice qualities found in various different styles of singing: Classical, Belt, Legit, R&B, Jazz, Country, and Pop. The data was elicited from a professional soprano and the voice qualities reflect industry stan-

dards. The data set for this paper is limited to samples using the vowel [i]. Laryngographic (LGG) data was generated simultaneously with the audio samples. This paper will focus on the results of the LGG analysis; however, an audio analysis was also performed using Spectrogram, LPC, and FFT. Data from the LGG is used to calculate the contact quotient, speed quotient, and ascending slope. The LGG waveform is also visually assessed. The LGG analysis gives insights into the source vibration for the different singing styles.

3aMU15. Optimized interpolations and nonlinearity in numerical studies of woodwind instruments. Apostolos Skouroupathis (Dept. of Phys., Univ. of Cyprus, P.O. Box 20537, Nicosia CY-1678, Cyprus, php4as01@ucy.ac.cy)

The impedance spectra of woodwind instruments with arbitrary axisymmetric geometry are studied. Piecewise interpolations of the instruments' profile are performed, using interpolating functions amenable to analytic solutions of the Webster equation. Our algorithm optimizes on the choice of such functions, while ensuring compatibility of wave-fronts at the joining points. Employing a standard mathematical model of a single-reed mouthpiece, as well as the time-domain reflection function which is derived from our impedance results, the Schumacher equation is solved for the pressure evolution in time. Analytic checks are made to verify that, despite the nonlinearity in the reed model and in the evolution equation, solutions are unique and singularity-free.

3aMU16. An alternative solution for positioning sound sources in loudspeaker-based virtual environments. Jonas Braasch (CIRMMT, McGill Univ., Montreal, QC, Canada H3A 1E3, jb@music.mcgill.ca)

In auditory virtual environments it is often required to position an anechoic point source in three-dimensional space. When sources in such applications are to be displayed using multichannel loudspeaker reproduction systems, the processing is typically based upon simple amplitude-panning laws. This paper describes an alternative approach based on an array of virtual microphones. The aim of the system architecture design was to comply with the expectations of audio engineers and to create sound imagery similar to those associated with standard sound recording practice. In the newly designed environment, the microphones, with adjustable directivity patterns and axis orientations, can be spatially placed as desired. This way, time-delay panning can be included by simulating a non-coincident microphone array. By the use of omnidirectional microphones the panning is almost solely based on time-delay panning and thus performance can take advantage of the precedence effect. Also, the proposed architecture is flexible enough to include classic amplitude panning laws by translating the requested panning law into physically non-existing microphone directivity patterns. Moreover, the virtual directivity patterns do not have to be rotationally symmetrical. [Work supported by VRQ.]

Session 3aNS**Noise: Progress in Managing Occupational Noise**

Hugh Davies, Cochair

Univ. of British Columbia, School of Occupational and Environmental Hygiene, 2206 East Mall, Vancouver, BC V6T 1Z3, Canada

Murray R. Hodgson, Cochair

*Univ. of British Columbia, School of Occupational and Environmental Hygiene, 2206 East Mall, Vancouver, BC V6T 1Z3, Canada***Chair's Introduction—8:00*****Invited Papers*****8:05**

3aNS1. Hearing conservation practices in eight noisy industries. William E. Daniell, Susan S. Swan, Janice Camp, Martin Cohen (Dept. of Environ. and Occupational Health Sci., Univ. of Washington, Box 357234, Seattle, WA 98195), Mary M. McDaniel (Pacific Hearing Conservation, Seattle, WA 98117), John Stebbins (Dept of Labor & Industries, Olympia, WA 98504), and Robert Leo (Dept of Labor & Industries, Seattle, WA 98104)

This study evaluated noise exposures and hearing conservation practices at 76 companies in eight industries with high rates of workers' compensation claims for hearing loss. Nearly all companies had exposures that required a hearing conservation program, and more than half had exposures that required consideration of noise controls. The use of noise measurements and consideration of controls was low in all industries. The completeness of hearing conservation programs was strongly associated with the extent of exposure in an industry, although practices varied widely within industries. Most companies had substantial deficiencies. More than one-third did not conduct annual training, and training had shortcomings at many others. One-third had not conducted audiometry. Hearing protection was commonly underused. Reported use was highest at companies with relatively complete programs, and in industries where exposure was most prevalent and least intermittent. Many employees had difficulty estimating how often, and presumably when, their exposure was excessive. There is a need for new strategies to promote and maintain hearing conservation efforts in noisy industries. The industries with greatest margin for improvement are not the noisiest industries but those where exposure is moderate or intermittent. [Work supported by the National Institute for Occupational Safety and Health.]

8:25

3aNS2. Effectiveness of the hearing conservation program: Change in hearing threshold shift incidence among industrial workers, 1978 to 2003. Hugh Davies (UBC School of Occupational and Environ. Hygiene, 2206 East Mall, Vancouver, BC, Canada V6T 1Z3, hugh.davies@ubc.ca)

Hearing conservation programs (HCP) are widely employed in preventing noise-induced hearing loss, but studies of their effectiveness have been rare. The impact of the implementation of hearing conservation programs was assessed in a large group of highly noise-exposed blue-collar workers by investigating time-trends in hearing-threshold shift incidence. Serial annual audiograms for employees of 14 British Columbia lumber mills for the period 1978 to 2003 were obtained from local regulatory-agency archives. Audiograms and concomitant otological medical histories were linked to subjects' work histories and noise exposure data. Multivariable Cox proportional hazard models were used to model the incidence of hearing threshold shift while controlling for age, baseline level of hearing loss, and other potential confounders. A total of 109 257 audiograms were associated with 10 590 subjects. Mean noise exposure in this group was 91.4 dBA(A). Mean interval between hearing tests was 566 days and mean age at first threshold shift was 44. Forty-six percent of subjects had at least one OSHA significant threshold shift during follow up. Preliminary analyses indicated a trend toward lower incidence of threshold shifts over the study period, with incidence in 5 approximately equal 5-year periods from 1978 to 2003 being 3.2%, 6.6%, 4.9%, 4.3% and 2.4%, respectively.

8:45

3aNS3. Use and effectiveness of hearing protection in construction. Richard Neitzel and Noah Seixas (Dept. of Env. and Occup. Health Sci., Univ. of Washington, 4225 Roosevelt Way NE #100, Seattle, WA 98105, rneitzel@u.washington.edu)

This paper presents an analysis of data on noise exposure and hearing protection device (HPD) use among construction workers drawn from several large datasets covering nine construction trades. A unique combination of one-minute dosimetry noise exposure level and simultaneous self-reported HPD use data were evaluated, as were occupational and non-occupational HPD use data collected by questionnaire as part of a longitudinal noise and hearing loss study among apprentices. Direct HPD attenuation measurements were also made on workers at their worksite. Workers reported using HPDs less than one-quarter of the time they were exposed above 85 dBA, the NIOSH Recommended Exposure Limit. Workers who reported always using HPDs in high noise were found to wear them only one-third of the time their exposures exceeded 85 dBA. Direct attenuation measurements indicated that workers achieved more

than 50% of the rated attenuation of their HPD on average, but that the variability in achieved attenuation was large. Lastly, when the measured HPD attenuation level and use time data were combined, the effective protection afforded by HPDs was less than 3 dB on average, though there was substantial variation among the different trades. These results demonstrate the need for better hearing conservation programs in construction.

9:05

3aNS4. Alternative noise exposure metrics among construction workers. Noah Seixas, Rick Neitzel, Lianne Sheppard, and Bryan Goldman (DEOHS, Univ. of Washington, Box 357234, Seattle, WA 98195-7234)

There continues to be some uncertainty in the choice of noise exposure metrics for assessing the risk of hearing damage. In particular, the Leq, based on a 3 dB exchange rate (ER) and the Lavg, which uses a 5 dB ER are both in use, as are various metrics representing exposure peaks. This paper presents an analysis of a large database of noise exposures among construction workers, comparing several noise metrics, and their application to a cohort of construction workers, followed audiometrically over a three year period. Metrics examined were the Lavg, Leq, and Lmax, expressing average levels. Two novel metrics were derived from these monitored metrics, Leq/Lavg, and Lmax/Leq, as measures of exposure variability and peakiness. A total of 730 workshifts, including data on 361 492 min of exposure to workers in 9 trades were examined. Correlations between average metrics (Leq, Lavg, Lmax) are generally very high, while the variability metrics are poorly correlated with average levels and with each other, indicating that they characterize different aspects of exposure. The task-specific estimates of individual exposure, using the various metrics, were applied to the cohorts work history and their effects on changes in hearing were assessed.

9:25

3aNS5. Special issues in hearing loss prevention in the Canadian military environment. Christian Giguère and Chantal Laroche (Univ. of Ottawa, 451 Smyth Rd., Ottawa, ON, Canada K1H 8M5)

Noise can be particularly noxious to hearing in the military. The personnel regularly face a wide range of noise-hazardous situations, many of which are seldom encountered in other work environments. High noise levels are associated with the operation of small arms and large caliber weapons, combat vehicles, aircrafts, ships and vessels, and industrial equipment. This can induce permanent and temporary hearing loss, compromise speech communication, interfere with the detection and localization of sound sources and warning sounds and thus, can jeopardize life or safety of the personnel. This paper will review the essential elements of a hearing loss prevention program proposed for the Canadian Armed Forces. The ultimate goal is to preserve hearing health as well as all hearing abilities necessary for effective operations. The program has been designed to meet the noise measurement and hazard investigation procedures, limits on noise exposure, use of hearing protection and other regulatory measures contained in the Canadian Occupational Health and Safety (COHS) Regulations (Part VII: Levels of Sound), while addressing the particular nature of the military environment. The paper will focus on issues that are not typically found in other occupational environments (variable work schedules, excessive impulse noise, extended exposures, communication devices).

9:45

3aNS6. In-situ personal assessment of hearing protector performance—recommendations for an updated standard. Jérémie Voix, Frédéric Laville (École de technologie supérieure, Université du Québec, Montreal (PQ), Canada H3C 1K3), and Jean Zeidan (Sonomax Hearing Healthcare Inc., Montreal (PQ), Canada H4P 2E2)

Standardized method currently used to assess Hearing Protection Devices (HPD) performance have little correspondence with protection achieved by users in practice. To overcome this problem new measurement method are being proposed. They no longer rely on statistically-based, population-based estimate, such as the Noise Reduction Rating (NRR), but on an individual assessment of the performance of the HPD. These new methods are highly beneficial for effective Hearing Conservation Programs as they provide individual attenuation values for noise-exposed workers in lieu of the statistically-based laboratory performance estimate, but their implementation will necessitate complete revision of existing standards. Among the issues to be addressed, are the aspect of the accuracy of the measurements performed by those methods and associated devices and also the replacement of any single number rating of HPD (such as the NRR) in existing standards. Solutions to address these issues and some general guidelines for an updated standard will be presented.

10:05–10:20 Break

10:20

3aNS7. Active hearing protectors: Prospects for improving speech intelligibility while maintaining noise reduction and audibility of warning alarms. Anthony J. Brammer, Donald R. Peterson, Martin G. Chorniack, and Subhash Gullapalli (Ergonomic Technol. Ctr., Univ. of Connecticut Health Ctr., 263 Farmington Ave., Farmington, CT 06030, anthonybrammer@hotmail.com)

Users of hearing protection devices (HPDs), including the communication headsets and helmets commonly worn by aircraft air and ground crews, emergency responders, military personnel, and persons in industry working in hostile environments, have long complained that HPDs interfere with their ability to communicate and to hear warning alarms. There thus appears to be a need for additional specialized HPDs designed specifically for these situations. The technology of active noise control together with signal processing applied to the communication channel of the HPD may provide a potential solution. The combination could improve noise reduction at low frequencies and enhance the intelligibility of speech reproduced by the communication channel, while maintaining the overall noise reduction within occupational exposure requirements, and assisting the perception of external warning alarms. A

strategy for achieving these goals using an active circumaural earmuff equipped with sub-band processing will be described, in which parallel controllers optimize the speech signal-to-noise ratio in bands that contribute most to intelligibility. The perception of a tonal alarm is assisted by binaural feed through to the earphones, within the constraints imposed by occupational exposure limits, and by sub-band processing to reduce the upward spread of masking.

Contributed Papers

10:40

3aNS8. The combined effect of noise and carbon monoxide on hearing thresholds of exposed workers. Adriana Lacerda, Tony Leroux, and Jean-Pierre Gagn (Univ. of Montreal, C.P.2168, succursale centre-ville, H3C 3J7, Montreal, Canada, Adriana.lacerda@umontreal.ca)

Animal models have been used to demonstrate the potentiation of noise-induced hearing loss (NIHL) by carbon monoxide. It has been shown that the addition of carbon monoxide to otherwise safe noise exposure levels produces significant NIHL in rats. However, the effects of chronic exposure to low level of carbon monoxide in a noisy work environment are still unknown. The aim of this study was to compare the hearing thresholds of a group of workers exposed to noise and carbon monoxide (Group 1) to another group of workers where carbon monoxide exposure is absent or negligible (Group 2). The analysis was based on 9396 audiograms collected by the Quebec National Public Health Institute between 1983 and 1996. The results show significantly poorer hearing thresholds at high frequencies (3, 4, and 6 kHz) for the carbon monoxide exposed group ($p < 0.001$). The potentiation effect also varied according to years of exposure in work place; a larger effect is observed for workers with between 15 to 20 years of exposure ($p < 0.001$). This study provides the first demonstration of a potentiation effect of NIHL by carbon monoxide in humans.

10:55

3aNS9. Derivation of an analytic expression for the error associated with the noise reduction rating. William J. Murphy (Hearing Loss Prevention Team, NIOSH 4676 Columbia Pkwy., MS C-27, Cincinnati, OH 45226-1998)

Hearing protection devices are assessed using the Real Ear Attenuation at Threshold (REAT) measurement procedure for the purpose of estimating the amount of noise reduction provided when worn by a subject. The rating number provided on the protector label is a function of the mean and standard deviation of the REAT results achieved by the test subjects. If a group of subjects have a large variance, then it follows that the certainty of the rating should be correspondingly lower. No estimate of the error of a protector's rating is given by existing standards or regulations. Propagation of errors was applied to the Noise Reduction Rating to develop an analytic expression for the hearing protector rating error term. Comparison of the analytic expression for the error to the standard deviation estimated from Monte Carlo simulation of subject attenuations yielded a linear relationship across several protector types and assumptions for the variance of the attenuations.

11:10–11:40

Panel Discussion

WEDNESDAY MORNING, 18 MAY 2005

REGENCY B, 8:30 A.M. TO 12:00 NOON

Session 3aPA

Physical Acoustics, Structural Acoustics and Vibration and Underwater Acoustics: Scattering From Non-Canonical Objects

Steven G. Kargl, Chair

Applied Physics Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698

Invited Papers

8:30

3aPA1. Intermediate scale measurement system for underwater acoustic measurements. Jan M. Niemiec, Paul C. Shang (NSWCCD, 9500 MacArthur Blvd., West Bethesda, MD 20817, niemiecjm@nswccd.navy.mil), and Stephan Schreppler (ONR, Code 334, Arlington, VA 22217)

The Intermediate Scale Measurement System (ISMS) provides high quality acoustic measurements on large-scale submarine models and other underwater structures in a controlled environment. Development and implementation of advanced marine technology designs require a thorough understanding of structural acoustic response mechanisms. The only practical and cost effective approach to developing this understanding is through the use of a combination of numerical models and large-scale physical models. The ISMS provides an affordable, unique capability of making precision acoustic measurements on models large enough to contain the structural detail and the instrumentation required to understand structural acoustic mechanisms and to confirm and guide the use of numerical models. This is a key element for prediction of full-scale performance and integration of advanced submarine technologies into operational systems.

3aPA2. Scattering by truncated targets with and without boundary interactions. Philip L. Marston, Kyungmin Baik, Aubrey Espana, Curtis F. Osterhoudt (Phys. Dept., Washington State Univ., Pullman, WA 99164-2814, marston@wsu.edu), Scot F. Morse (Western Oregon Univ., Monmouth, OR 97361), Brian T. Hefner (Univ. of Washington, Seattle, WA 98105), and Florian J. Blonigen (Northeastern Univ., Boston, MA 02115)

Ray methods have been applied to the scattering of various truncated targets having wavenumber-radius products as small as 10 [F. J. Blonigen and P. L. Marston, *J. Acoust. Soc. Am.* **107**, 689–698 (2000); S. F. Morse and P. L. Marston, *ibid.* **112**, 1318–1326 (2002); B. T. Hefner and P. L. Marston, *ARLO* **2**, 55–60 (2001)]. Recent work emphasizes the exploration of scattering enhancements for other situations including plastic cylinders having curved ends, truncated plastic cones, partially exposed cylinders, and objects in simulated conditions for burial in a seabed. Enhanced scattering is often associated with a locally flat outgoing wavefront. For plastic targets it has been helpful to examine the time dependence of the backscattered envelope as a function of target tilt for targets illuminated by short tone bursts. For partially exposed objects it is helpful to examine the backscattering as a function of the target exposure. For simulated buried targets, it has been helpful to excite target resonances. [Work supported by the Office of Naval Research.]

9:20

3aPA3. Finite-element modeling of acoustic scattering from realistic elastic structures. David S. Burnett (Naval Surface Warfare Ctr., Code R21, 110 Vernon Ave., Panama City, FL 32407, david.s.burnett@navy.mil)

Over the last twenty years the author has developed several finite-element codes for modeling acoustic scattering from complicated elastic structures. All the codes have employed 3-D continuum mechanics throughout the computational domain. Thin structural components, such as plates and shells, are modeled with 3-D elasticity rather than plate or shell theories. The codes are therefore fully 3-D, i.e., in both physics and geometry. The talk will first explain the rationale for this approach and touch briefly on essential aspects of previous codes. It will then focus in more detail on the current code under development, which employs relatively inexpensive commercial software well-suited for this fully 3-D approach.

Contributed Papers

9:45

3aPA4. Acoustic scattering from an obstacle in a homogeneous non-uniform moving medium via the ray method. Tatiana Semenova (Natl. Res. Council Postdoctoral Res. Associate, NASA Langley Res. Ctr., Hampton, VA 23666)

For an accurate noise prediction for aircraft in forward flight, scattering of engine noise from airframe must be included into the description of the sound field. A realistic model of acoustic scattering by airframe in flight must account for the effects of the background flow. We consider acoustic scattering from an obstacle in the presence of a time-independent background flow at low Mach number and we neglect all other effects of the variation of the medium on sound propagation. The model is based on a convected Helmholtz equation with coefficients dependent on the background flow, which can be given either analytically or by CFD. A high-frequency solution is sought by the ray acoustics with a simple method based on the Fresnel volumes of the rays employed to determine the caustic zones and the field on the caustics. The weaker effects of a tip or smooth-body diffraction are neglected and only the reflection and edge diffraction are taken into account. The source of noise is modeled as a point source. The upstream shielding effect of a boundary layer is investigated as an example and uniform versus non-uniform flow effects are illustrated. Other numerical examples will be considered. [Work supported by NRC.]

10:00–10:30 Break

10:30

3aPA5. Scattering by tilted plastic cylinders having curved ends and truncated plastic cones. Aubrey Espana, Kyungmin Baik, and Philip L. Marston (Phys. Dept., Washington State Univ., Pullman, WA 99164-2814, marston@wsu.edu)

In prior research an acoustic backscattering enhancement was demonstrated for a bluntly truncated plastic cylinder caused by a merged caustic [F. J. Blonigen and P. L. Marston, *J. Acoust. Soc. Am.* **107**, 689–698 (2000)]. This was confirmed with analogous light scattering experiments [P. L. Marston, Y. B. Zhang, and D. B. Thiessen, *Appl. Opt.* **42**, 412–417 (2003)]. In recent work a different backscattering enhancement associated with a caustic was identified for tilted plastic cylinders having curved

ends. When the cylinder is tilted so as to focus a shear wave at the point of internal specular reflection, the curvature of the outgoing acoustic wavefront vanishes orthogonal to the meridional plane. This was verified with analogous light scattering experiments. The flatness of the outgoing wavefront enhances the scattering. Backscattering by truncated plastic cones as a function of tilt also shows enhancements associated with the composition of the target. The time dependence of the backscattering envelope as a function of tilt reveals different features depending on whether the top or bottom of the cone is illuminated by tone bursts. [Work supported by the Office of Naval Research.]

10:45

3aPA6. Effect of changes of complex surface impedance for acoustic scattering from a wedge. Ron G. Hughes, Jan M. Njemić (NSWC Carderock Div., W. Bethesda, MD 20817), and Herbert Überall (NSWC Carderock Div., W. Bethesda, MD 20817)

A code written earlier for the calculation of acoustic scattering from a submerged wedge has been employed in order to study the dependence of the scattering amplitude on various assumed surface impedances of the wedge. Typical wedge opening angles are 10° . Plane waves are assumed incident with source and observation points coinciding in the far field along the bisector of the wedge. The target strength for an impedance-coated wedge is found to lie in between those of a rigid and a pressure release wedge. A frequency region up to 20 kHz was considered. Various numerical impedances and angles of incidence are assumed.

11:00

3aPA7. Analysis of scattering by spheres having a negative acoustical refractive index. Philip L. Marston (Phys. Dept., Washington State Univ., Pullman, WA 99164-2814, marston@wsu.edu)

Electromagnetic waves having oppositely directed phase and group velocities propagate in metamaterials having a negative permeability and negative permittivity [J. B. Pendry and D. R. Smith, *Phys. Today* **57**(6), 37–44 (2004)]. Such materials are predicted to have unusual electromagnetic scattering properties [R. Ruppin, *Solid State Commun.* **116**, 411–415 (2000)]. If it is possible to fabricate acoustical materials having a simultaneously negative effective elastic modulus and density (in a dynamical sense), the mechanical energy flux will have the opposite direction as the

wave-vector associated with phase evolution. Rays descriptive of the energy flux refracted by such hypothetical materials at interfaces with ordinary fluids would be characterized by a negative acoustical refractive index. Partial-wave-series calculations of high frequency scattering by fluid spheres having an acoustical refractive index at (or close to) 1 reveal backscattering enhancements associated with glory rays which, unlike ordinary spheres [P. L. Marston and D. S. Langley, *J. Acoust. Soc. Am.* **73**, 1464–1475 (1983)], require only a single internal chord. Generalized Lamb waves on elastic shells having opposite phase and group velocities also cause enhanced backscattering associated with unusual rays [G. Kaduchak, D. H. Hughes, and P. L. Marston, *J. Acoust. Soc. Am.* **96**, 3704–3714 (1994)].

11:15

3aPA8. Production of evanescent acoustic waves and their scattering by resonant targets. Curtis F. Osterhoudt, Christopher Dudley, David B. Thiessen, Philip L. Marston (Dept. of Phys. and Astron., Washington State Univ., Pullman, WA 99164-2814, gardyloo@mail.wsu.edu), and Scot F. Morse (Western Oregon Univ., Monmouth, OR 97361)

An environmentally-friendly liquid mixture, denser than water and immiscible with it, and with a sound speed slower than water's, has been produced and incorporated into a scattering facility. This liquid mixture is submerged in a volume of water. In this way, the water above the denser fluid simulates a sandy ocean bottom, with the dense fluid simulating the ocean water column. When insonified by an acoustic beam from within the denser liquid, evanescent wave fields are produced in the water volume. Various targets and detectors are scanned in the water (the simulated sandy bottom) or in the mixture (the simulated ocean water column), and coupling of the evanescent waves to low-frequency resonances of cylindrical targets has been observed. Such coupling may reveal important structural characteristics of scatterers. Additionally, the complicated soundfield structure in the simulated bottom has been mapped. The structure is associated with the production of evanescent acoustic waves by a beam of sound. A model for this structure was developed using numerical wave-number integration. [Work supported by the Office of Naval Research.]

11:30

3aPA9. Sound scattering by clusters of rigid cylinders in air. Miguel Figueres, Francisco Cervera, Andreas Håkansson, José Sánchez-Dehesa, and Jaime Llinares (Univ. Politécnic de Valencia, E-46022 Valencia, Spain, jsdehesa@upvnet.upv.es)

The scattering of sound by a cluster of rigid cylinders in air has been investigated both theoretically and experimentally. The cluster was made by a lattice of cylinders with hexagonal symmetry. The sound scattered in all directions around the cluster is measured and compared with a calculation based on multiple scattering theory. The behavior as a function of the external geometry of the cluster will also be presented. The refraction and reflection of the sound by the boundaries of the cluster is analyzed by a theoretical model that takes into account its inner crystal symmetry. [Work supported by MEC of Spain.]

11:45

3aPA10. Complete regularization of boundary integral equations in wave field diffraction problems on curved surfaces. German A. Maximov and Dmitry N. Lesonen (Dpt. 39, Moscow Eng. Phys. Inst., Kashirskoe sh. 31, Moscow, 115409, Russia, maximov@dpt39.mephi.ru)

The method of regularization of exact integral Fredholm equations of the second kind for field or its normal derivative on the scattering surface was suggested in the previous work. It allows to obtain the stable solutions, including ones for the resonant domains, when the direct numerical solution of exact initial integral equations results to unstable solutions. Mathematically the regularization concludes in the splitting of exact integral Fredholm equations on the equivalent system of two equations, one of them possesses the restricted by geometrically illuminated area integral operator, and other one corresponds by form to an initial one, but with diffracted source. The solution of the first equation is considered as regularized one, which coincides with exact solution in the short wavelength limit. From the physical point of view the regularization of initial integral equations concludes in their replacement on the restricted analogs, where the contributions of geometrically shadowed areas are eliminated. In the given work the regularization of the rested after splitting integral equation is represented and it is shown that the process of such regularization is auto-similar one, so that it is possible to carry out the complete regularization of the problem.

WEDNESDAY MORNING, 18 MAY 2005

REGENCY C, 8:30 TO 11:50 A.M.

Session 3aPP

Psychological and Physiological Acoustics: Hearing With More Than One Ear

Richard L. Freyman, Chair

Dept. of Communicative Disorders, Univ. of Massachusetts, 125 Arnold House, Amherst, MA 01003

Chair's Introduction—8:30

Contributed Papers

8:35

3aPP1. High resolution auditory perception system. Iftekhar Alam (1013-Clover Heights, Opp Salunke Vihar, Salunke Vihar Rd., Wanowrie, Pune-411040, India [iftekharamk@yahoo.com]) and Ashok Ghatol (Pune Inst. of Engg. and Technol. Shivajinagar, Pune, Maharashtra, India)

Blindness is a sensory disability which is difficult to treat but can to some extent be helped by artificial aids. The paper describes the design aspects of a high resolution auditory perception system, which is designed on the principle of air sonar with binaural perception. This system is a

vision substitution aid for enabling blind persons. The blind person wears ultrasonic eyeglasses which has ultrasonic sensor array embedded on it. The system has been designed to operate in multiresolution modes. The ultrasonic sound from the transmitter array is reflected back by the objects, falling in the beam of the array and is received. The received signal is converted to a sound signal, which is presented stereophonically for auditory perception. A detailed study has been done as the background work required for the system implementation; the appropriate range analysis procedure, analysis of space-time signals, the acoustic sensors study, amplification methods and study of the removal of noise using filters. Finally the system implementation including both the hardware and the software

part of it has been described. Experimental results on actual blind subjects and inferences obtained during the study have also been included.

8:50

3aPP2. Sex differences present in auditory looming perception, absent in auditory recession. John G. Neuhoff (Dept. of Psych., The College of Wooster, Wooster, OH 44691) and Erich Seifritz (Univ. of Bern Switzerland)

When predicting the arrival time of an approaching sound source, listeners typically exhibit an anticipatory bias that affords a margin of safety in dealing with looming objects. The looming bias has been demonstrated behaviorally in the laboratory and in the field (Neuhoff 1998, 2001), neurally in fMRI studies (Seifritz *et al.*, 2002), and comparatively in non-human primates (Ghazanfar, Neuhoff, and Logothetis, 2002). In the current work, male and female listeners were presented with three-dimensional looming sound sources and asked to press a button when the source was at the point of closest approach. Females exhibited a significantly greater anticipatory bias than males. Next, listeners were presented with sounds that either approached or receded and then stopped at three different terminal distances. Consistent with the time-to-arrival judgments, female terminal distance judgments for looming sources were significantly closer than male judgments. However, there was no difference between male and female terminal distance judgments for receding sounds. Taken together with the converging behavioral, neural, and comparative evidence, the current results illustrate the environmental salience of looming sounds and suggest that the anticipatory bias for auditory looming may have been shaped by evolution to provide a selective advantage in dealing with looming objects.

9:05

3aPP3. Lateralization of high-frequency transposed stimuli under conditions of binaural interference. Leslie R. Bernstein and Constantine Trahiotis (Dept. of Neurosci. and Dept. of Surgery (Otolaryngol.), Univ. of Connecticut Health Ctr., Farmington, CT 06030, Les@neuron.uconn.edu)

The purpose of this study was to determine whether binaural interference would occur if ITD-based extents of laterality were measured using high-frequency transposed stimuli as targets. The results of an earlier study [L. R. Bernstein and C. Trahiotis, *J. Acoust. Soc. Am.* **116**, 3062–3069 (2004)], which focused on threshold-ITDs rather than extents of laterality, suggested that high-frequency transposed stimuli might be immune to binaural interference effects resulting from the addition of a spectrally-remote, low-frequency interferer. In contrast to the earlier findings, the data from this study indicate that high-frequency transposed targets can, indeed, be susceptible to binaural interference. High-frequency transposed targets, even when presented along with an interferer, yielded greater extents of ITD-based laterality than did Gaussian noise targets presented in isolation. That is, the enhanced potency of ITDs conveyed by transposed stimuli persisted even in the presence of a low-frequency interferer. Predictions made using an extension of the model of Heller and Trahiotis [L. M. Heller and C. Trahiotis, *J. Acoust. Soc. Am.* **99**, 3632–3637 (1996)] accounted well for binaural interference obtained with conventional Gaussian noise targets but generally over-predicted the amounts of interference found with the transposed targets.

9:20

3aPP4. Binaural models for the detection of interaural incoherence. Matthew J. Goupell (Dept. of Phys. and Astron., Michigan State Univ., 4230 BPS Bldg., East Lansing, MI 48824) and William M. Hartmann (Michigan State Univ., East Lansing, MI 48824)

Human listeners have a remarkable ability to detect small amounts of interaural incoherence in bands of noise. Experiments using frozen noise, all with the same value of interaural coherence (0.9922), show that when the bandwidth is large the ability to detect incoherence is essentially determined by the value of coherence itself. However, when the bandwidth is narrow, e.g. 20 Hz, detection depends on some function of the moment-

to-moment fluctuations in interaural phase differences and interaural level differences. Experiments using hundreds of different frozen noises and two bandwidths were performed to find the best binaural model for predicting detection performance for such narrow bands. Successful models incorporated known elements of binaural hearing—temporal averaging and compression. Models with independent processing of interaural phase fluctuations and interaural level fluctuations were marginally more successful than models based on fluctuations in lateral position. The most successful models correlated well with average human responses, $r = 0.89$. It seems likely that such individual fluctuations remain dominant in masking level difference experiments in the narrow band limit. [Work supported by the NIDCD grant DC 00181.]

9:35

3aPP5. A binary masking technique for isolating energetic masking in speech perception. Douglas S. Brungart, Brian D. Simpson (AFRL/HECB, 2610 Seventh St., WPAFB, OH 45433, douglas.brungart@wpafb.af.mil), Peter S. Chang, and DeLiang Wang (Ohio State Univ., Columbus, OH 43210)

When a target speech signal is obscured by interfering speech, two distinct types of masking contribute to the resulting degradation in the intelligibility of the target talker: energetic masking caused by overlap in the time-frequency distribution of energy in the two voices, and informational masking caused by the listener's inability to correctly segregate the acoustic elements of the two voices into distinct streams. This study attempted to isolate the effects of energetic masking on multitalker speech perception with ideal time-frequency binary masks that retained those spectro-temporal regions of the acoustic mixture that were dominated by the target speech but eliminated those regions that were dominated by the interfering speech. This procedure removed the same phonetic information from the target speech that would ordinarily be lost due to energetic masking, but eliminated the possibility for the kinds of target-masker confusions that are thought to produce informational masking. The results suggest that energetic masking may play a surprisingly small role in the overall masking that occurs in certain types of multitalker speech signals. They also indicate that the number of competing talkers has a much greater influence than target-masker similarity on the amount of energetic masking that occurs in a multitalker stimulus.

9:50

3aPP6. Binaural effects in simultaneous room reflection masking. Joerg M. Buchholz (MARCS Auditory Labs., Univ. of Western Sydney, Locked Bag 1797, Penrith South DC NSW 1797, Australia)

Masked thresholds (MT) for a single test reflection masked by a direct sound (200 ms long broadband noise) were measured dependent on the time delay of the reflection for diotic as well as dichotic stimulus presentation. In the diotic case, the direct sound and the test reflection were presented equally to both ears via headphones. In the dichotic case, an ITD of 0.5 ms was added to the test reflection. In order to focus on simultaneous masking effects, the reflection was truncated in such a way that it formed a common offset with the direct sound. For the diotic case, the resulting data showed a MT increase with increasing reflection delay and for the dichotic case a MT decrease with increasing reflection delay, producing an intercept between both curves at a reflection delay of 6–8 ms. Hence, negative BMLDs (up to –8 dB) were found for very early reflections and positive BMLDs (up to +8 dB) for later reflections, suggesting a binaural mechanism that suppresses very early reflections and enhances later reflections. The measurement results are discussed in the background of different auditory models.

10:20

3aPP7. The influence of target-masker similarity in a three-talker diotic listening task. Nandini Iyer and Douglas Brungart (Air Force Res. Labs, 2610 Seventh St., Wright–Patterson AFB, OH 45433)

In diotic cocktail-party listening tasks, performance decreases dramatically when the number of talkers increases from two (one target and one masker) to three (one target and two maskers). In this experiment, listeners were asked to extract information from a Coordinate Response Measure (CRM) target phrase that was masked by one or two interfering utterances. The maskers were either semantically similar to the target (i.e., CRM phrases with different call signs than the target phrase) or semantically different from the target (time-reversed, contextually dissimilar, or in a different language than the target phrase). The two-talker results showed that performance was much worse when the interfering phrase was semantically similar to the target than when it was semantically different. However, given that one interfering talker was similar to the target, the contextual similarity of the second interfering talker had very little effect on performance; in other words, target intelligibility in the presence of two CRM maskers was no worse than with one CRM masker and one time-reversed or foreign-language masker. These results suggest that the disruption of speech segregation caused by the addition of a second interfering talker cannot be explained solely by contextual similarity between the target and masking speech.

10:35

3aPP8. Onsets, autocorrelation functions and spikes for direction-based sound source separation. Leslie S. Smith and Dagmar S. Fraser (Dept. of Computing Sci. and Mathematics, Univ. of Stirling, Stirling FK9 4LA, Scotland, lss@cs.stir.ac.uk)

The aim of this work is separation of foreground speech from background sound sources using selective remixing of bandpass filtered channels. Clearly, the remixing parameters must be dynamic since the speech and noise spectra are highly non-stationary. Remixing parameters are re-computed at onsets, detected using biologically motivated techniques [L. S. Smith and D. S. Fraser, *IEEE TNS* **15**, 1125–1134 (2004)]. However, onsets may originate from the foreground or the background. To select appropriate onsets from the foreground source (whose direction is known) a two microphone system is used, selecting onsets for which the estimated direction in that channel corresponds to the foreground direction. Two different techniques for direction estimation are used: a channel by channel short-term autocorrelation technique, and a channel by channel spike based phase synchronous system (SBPSS), computing ITDs [L. S. Smith, in *Artificial Neural Networks*, Proc ICANN 2001, LNCS 2130, pp. 1103–1108 (Springer, 2001)] and IIDs [L. S. Smith, in *From Animals to Animats*, Vol. 7, pp. 60–61 (MIT Press, 2002)]. Results comparing the performance of autocorrelation and SBPSS on single source and source plus noise signals in an office environment are presented. [Work supported by UK EPSRC.]

10:50

3aPP9. How does the mammalian brain code for binaural information? Torsten Marquardt and David McAlpine (Dept. Physiol., UCL, Gower St., London WC1E 6BT, UK)

We will summarize new physiological results which contradict established models of binaural hearing, namely (1) the A-limited range of encoded interaural time difference, ITD, (2) a correlation between ipsilateral inhibition and best ITD, and (3) a frequency dependent non-linear mapping of ITD which reflects distribution of naturally occurring ITDs. Furthermore, we will outline a new concept for binaural models which incorporates these new results, and will propose a new illustrative format in which to present the triplet of binaural parameters: interaural phase difference, interaural intensity difference, and interaural coherence (redefined).

3aPP10. Adjustment of interaural time difference in head related transfer functions based on listeners' anthropometry and its effect on sound localization. Yōiti Suzuki, Kanji Watanabe, Yukio Iwaya (R.I.E.C./G.S.I.S., Tohoku Univ., 2-1-1 Katahira, Aoba-ku, Sendai, 980-8577, Japan), Jiro Gyoba (Tohoku Univ., Aoba-ku, Sendai, 980-8576, Japan), and Shouchi Takane (Akita Pref. Univ., Tsuchiya, Honjo, Akita, 015-0055, Japan)

Because the transfer functions governing subjective sound localization (HRTFs) show strong individuality, sound localization systems based on synthesis of HRTFs require suitable HRTFs for individual listeners. However, it is impractical to obtain HRTFs for all listeners based on measurements. Improving sound localization by adjusting non-individualized HRTFs to a specific listener based on that listener's anthropometry might be a practical method. This study first developed a new method to estimate interaural time differences (ITDs) using HRTFs. Then correlations between ITDs and anthropometric parameters were analyzed using the canonical correlation method. Results indicated that parameters relating to head size, and shoulder and ear positions are significant. Consequently, it was attempted to express ITDs based on listener's anthropometric data. In this process, the change of ITDs as a function of azimuth angle was parameterized as a sum of sine functions. Then the parameters were analyzed using multiple regression analysis, in which the anthropometric parameters were used as explanatory variables. The predicted or individualized ITDs were installed in the nonindividualized HRTFs to evaluate sound localization performance. Results showed that individualization of ITDs improved horizontal sound localization.

11:20

3aPP11. Spatial perception of motion-tracked binaural sound. Joshua B. Melick, V. Ralph Algazi, and Richard O. Duda (CIPIC Interface Lab., UC Davis, Davis, CA 95616)

Motion-tracked binaural sound reproduction extends conventional headphone-based binaural techniques by providing the dynamic cues to sound localization produced by voluntary head motion [V. R. Algazi, R. O. Duda, and D. M. Thompson, *J. Aud. Eng. Soc.* **52**, 1142–1156 (2004)]. It does this by using several microphones to sample the acoustic field around a dummy head, interpolating between the microphone signals in accordance with the dynamically measured orientation of the listener's head. Although the provision of dynamic cues reduces the sensitivity of the method to characteristics of the individual listener, differences between the scattered field produced by the dummy head and the scattered field that would be produced by a particular listener distorts the spatial perception. A common observation is that sound sources appear to rise in elevation when the listener turns to face them. We investigate this effect by comparing the perceived rise in elevation under three different conditions: recordings in which recordings are made using (a) the listener's own head, (b) a KE-MAR mannequin, and (c) a cylindrical head with no torso. Quantitative results are presented showing the degree to which perceptual distortions are least for (a) and greatest for (c). [Work supported by NSF.]

11:35

3aPP12. Transaural virtual reality applied to front-back localization. Peter Xinya Zhang (Dept. of Phys. and Astron., Michigan State Univ., 4230 BPS Bldg., East Lansing, MI 48824) and William M. Hartmann (Michigan State Univ.)

Human sound localization can be studied by measuring head-related transfer functions for different source locations and subsequently using the data to simulate different spatial locations using headphone presentation. This technique fails to adequately simulate the difference between front and back locations because the frequencies relevant to pinna cues are so high. A transaural (cross-talk cancellation) technique has been developed that is accurate up to 16 kHz. By controlling the levels and phases of spectral components, as measured by probe microphones in the ear canals, the technique eliminates significant differences between the baseline simulation and reality. The technique has been used, with modifications to the baseline simulation, to study the front-back information content in differ-

ent frequency bands by eliminating the information while maintaining the energy. It has also been used to test for orthogonality between front-back localization and azimuthal localization based on simple interaural differences. The technique has verified that front-back discrimination cannot be

mediated by interaural differences alone. By flattening the level spectrum in one ear while leaving the other ear unchanged it has been possible to demonstrate an ear advantage for front-back discrimination. [Work supported by the NIDCD, grant DC 00181.]

WEDNESDAY MORNING, 18 MAY 2005

PLAZA B, 8:30 A.M. TO 12:00 NOON

Session 3aSA

Structural Acoustics and Vibration: Approximation Techniques for Fluid Loading Effects I

David Feit, Cochair

Applied Physical Sciences, Inc., 2 State St., New London, CT 06320

Joel M. Garrelick, Cochair

Cambridge Acoustical Associates/ETC, 84 Sherman St., Cambridge, MA 02140

Invited Papers

8:30

3aSA1. The rational function approximation for fluid loaded vibrating structures. David Feit (Appl. Physical Sci. Corp., Ste. 300, 2 State St., New London, CT 06320)

A complicating feature of fluid-structure interaction problems is the non-local or long-range effect of the fluid loading acting on a structure vibrating in response to time-harmonic loading while submerged in an acoustic medium. DiPerna and Feit [J. Acoust. Soc. Am. **114**, 194–199 (2003)] recently introduced an approximation for this effect using a rational function approximation (RFA) to the spectral impedance. The latter relates the transform of the surface pressure to that of the surface velocity, and transforming back to the physical domain gives rise to a high-order differential pressure-velocity relationship. The coefficients of the n th order differentials are inversely proportional to the n th order of frequency so that the approximation gets better with increasing frequency. The approximation originally developed for a planar vibrating surface, has now been extended to a cylindrical surface, and this development and several applications of the RFA technique are presented. [Work supported by the Carderock Division Naval Surface Warfare Center Intra-laboratory Independent Research (ILIR) program and the Office of Naval Research.]

9:00

3aSA2. Transmission line analysis of cylindrical shells with exterior fluid loading. Ann Stokes, Charles Corrado, and Joel Garrelick (Appl. Physical Sci., Inc., 2 State St., Ste. 300, New London, CT 06320)

The Wave Transmission Line (WTL) numerical modeling formulation has been developed to understand and efficiently compute the transmission of acoustic energy through fluid-filled elastic shells. The structural systems are represented as a sequence of components and discontinuities connected with straight duct sections. The motion of straight segments is described in terms of amplitudes of freely traveling, left-and-right propagating waves. Recent work in Rational Function Approximation (RFA) of exterior fluid loading permits this approach to be extended to cylindrical structures with interior and exterior fluid loading. The resulting approach is more computationally efficient than conventional finite element approaches. This paper describes development of systems of transmission line equations for segmented cylindrical shells with exterior and interior fluid loading.

9:30

3aSA3. The Fresnel-Product Approximation, and comments on extending rational function approximation to include flow. Rudolph Martinez and Carina Ting (Anteon Corp., 84 Sherman St., Cambridge, MA 02140)

The first half of this presentation reviews the existing development of the Fresnel-Product Approximation as a replacement of the exact fluid impedance on a two-sided blade-like object. Excellent performance versus exact calculations is demonstrated for vibration patterns above, below, and at coincidence, viz., for flexural wavelengths that are longer, shorter, or equal to that of characteristic sound. The approximate solution is global and thus of the WKB type: (1) It captures the correct edge behavior of diffraction loads on the vibrating blade; (2) Collapses to the Kirchhoff solution away from both edges at high frequencies; and (3) Differs by only a small factor when applied at low frequencies to predict the blades globally entrained fluid mass. The second half of the talk addresses the Rational Function Approximation (RFA) pioneered by Feit and DiPerna. It shows how their spectral technique could be readily generalized to add flow effects, including the influence of shed vorticity as it co-exists with sound. Flowless acoustics will be derived from fully unsteady aerodynamics/aeroacoustics through a delicate limiting process and an integration by parts. The talk ends with a discussion of how current applications of the RFA could be stripped of some of their structural-acoustic artifacts.

10:15

3aSA4. Rational approximations of surface impedance operators on curved surfaces with application to high-frequency scattering by thin shells. David C. Calvo (Acoust. Div. Naval Res. Lab., Washington, DC 20375-5320)

In the analytical solution of scattering and radiation problems on curved surfaces using Fourier transforms, surface impedance functions featuring ratios of Hankel functions arise. Recently, an approach of approximating the impedance by a rational function was introduced in which the squared transform variable is replaced by a differential operator. By taking into account the impedance function poles in the rational approximation, an accurate and efficient numerical scheme is obtained applicable to wide-frequency ranges and wide-scattering angles where creeping waves are important. The rational approximation and operator introduction in effect creates an inverse Fourier transform operator so explicit residue calculus inversion is not needed if numerical results are desired. For general convex surfaces, the approximate impedance operator is a wide-angle on-surface radiation condition operator with more accuracy than the narrow-angle one of Kriegsmann *et al.* The use of rational approximations resembles that of Collins *et al.* for the square-root function in ocean acoustics and DiPerna & Feit for the flat plate, the important difference being the curvature. Results of using this approach for two and three-dimensional scattering problems will be presented along with recent application to scattering by thin cylindrical and spherical shells at high frequencies. [Work supported by ONR.]

Contributed Papers

10:45

3aSA5. Well- and ill-considered numerical solution constructions. John J. McCoy (The Catholic Univ. of America, Washington, DC 20064)

One requirement for a prediction problem to be well posed is that its solution must not be unduly sensitive to small changes in the problem specification. A similar requirement can be demanded for the numerical construction of the solution of a well-posed problem to be well considered; the achieved solution must not be unduly sensitive to small numerical errors that are unavoidable in the construction's implementation. Ill-considered constructions typically result on reformulating a well-posed problem, now as a series of sub problems, one of which is ill posed. Three classes of potentially ill-considered constructions of the steady state vibration field in a bounded continuous structure and of the narrow band acoustic field in an extended propagation environment appear prominently in the literature. One class is solution constructions based on "marching." A second class is solution constructions based on "substructuring" the computational domain. A third class is solution constructions that are based on a Born-series representation. Examples of well- and ill-constructions for each of the classes are briefly described, with well- and ill-considered marching solutions discussed in some detail.

11:00

3aSA6. Elastoacoustic response of laminated plates. Ramesh Kolar (Naval Postgrad. School, Mech. & Astro Eng., Code ME/Ko, Monterey, CA 93943)

The application of composite materials in the aerospace and naval structures has increased enormously due to high specific strength and specific stiffness afforded by these materials. In this paper a formulation is developed based on Hamilton's Principle and laminated composite plate theory to study the elasto-acoustical response of composite plates under heavy fluid loadings. The formulation starts by using Hamilton's principle in conjunction with shear deformable theory of laminated composite plates. The acoustic pressure described by wave equation is computed similar to Sandman and Nelisse. Using the Rayleigh Ritz method and symbolic mathematics for evaluation of integrals, the formulation provides efficient approach for the problem defined. Typical results include radiation impedance as a function of driving frequency, vibroacoustic indicators such as radiated sound power and mean square velocity for a model problem. Such results are very important in studying constrained layer damping when viscoelastic dampers are used in structural components.

11:15

3aSA7. Prediction of radiation from finite-length fluid-loaded periodic structures subject to broadband excitation using local-global homogenization. Pavel V. Danilov and Donald B. Bliss (Mech. Eng. and Mater. Sci., Duke University, Durham, NC 27708, dbb@duke.edu)

Local-Global Homogenization (LGH) is a method to predict directly the smooth global response of periodic fluid-loaded structures in a self-contained manner. Many fluid-loaded structures, such as fuselages and hulls, have periodically spaced braces, ribs, or attachments. Structural motion, acoustic radiation, and the interior sound field are of interest. Calculating the motion of such fluid-loaded structures is difficult because of their complexity and the disparity of length scales. Periodic discontinuities cause the structural response to occur in a broad spectrum of spatial wavenumbers, and to exhibit stop-band and pass-band behavior. The broad spatial wavenumber spectrum contains both radiating and non-radiating components. The global low-wavenumber part of the response is most efficiently coupled to the acoustic field, since low wavenumbers correspond to supersonic phase speeds. In the LGH reformulation, an infinite order operator that embodies both the structural modes and the evanescent component of the fluid loading governs the equivalent smooth global problem. Numerical implementation is demonstrated for the structural response and acoustic radiation from a finite-length fluid-loaded plate with periodic impedance discontinuities subject to broadband excitation. Calculations show good agreement with the exact solution and substantially improved computational efficiency. Radiation characteristics are discussed for single frequency and broadband structural motions.

11:30

3aSA8. Vibroacoustic optimization of skin damped structures for noise reduction. Lionel Zoghaib (Lab. de Mécanique et d'Acoustique CNRS, Marseille, France, zoghaib@lma.cnrs.mrs.fr)

Most of the time the widely used passive treatments are designed to obtain the highest achievable damping. Although this might be satisfactory from a vibrational point of view the approach is not optimal for noise reduction. The idea is illustrated in the case study of a plate on which polymer/aluminum thin patches have been bonded. The system acoustic pressure and structure displacement equations are transformed into a single integro differential one governing the plate displacement. The solution is then expanded into a series of resonance modes. These modes are estimated using a perturbation expansion with respect to the fluid/solid density ratio. Numerically, in vacuo resonance modes are firstly computed using a viscoelastic incompressible 3D FE program together with a Newton-type iterative method and a sparse complex matrix solver implementing ARPACK. Radiation modal impedances are then computed to ponderate the in vacuo solution. The numerical model has been experimentally validated. Its specificity lies in the separate computation of three damping components: thermoelastic, skin and acoustic damping, this last

one being caused by the fluid/structure coupling. Finally an optimization process determines an optimal system configuration to achieve the lowest acoustic damping while obtaining the highest viscoelastic damping.

11:45

3aSA9. An impedance modeling technique of a fluid-loaded structure. Chih-Chun Cheng and Pi-Wen Wang (Dept. of Mech. Eng., Natl. Chung Cheng Univ., 160, San-Hsing, Ming-Hsiung, Chia-Yi 621, Taiwan)

A methodology of modeling the fluid-loaded structures using the impedance technique is proposed. The structural response and the fluid loading effect are expressed in terms of structural impedance and the acoustic wave impedance. Then a formulation that assembles the acoustic imped-

ance and the structural impedance is derived and can be used to determine the response of a fluid-loaded structure. The advantage is simply that the impedance can be obtained either experimentally or analytically and then the response of the fluid-loaded structure can be found using the proposed impedance coupling formula without deriving the equation of motion which usually encounters difficulties when the structure is in an irregular shape or the associated boundary conditions are complicated. Two numerical examples are presented. The first is to validate that the fluid loading calculated using the proposed methodology is the same as that from the traditional Fourier Transformation method. The second is to demonstrate the potential of this proposed method applied in analyzing an active material system which is commonly used in the vibro-acoustic control. [Work supported by NSC of Taiwan.]

WEDNESDAY MORNING, 18 MAY 2005

REGENCY D, 8:00 A.M. TO 12:00 NOON

Session 3aSC

Speech Communication: Indigenous Languages of North America (Lecture/Poster Session)

Sonya F. Bird, Chair

Dept. of Linguistics, Univ. of Victoria, Victoria, BC V8W 3P4, Canada

Chair's Introduction—8:00

Invited Papers

8:05

3aSC1. Phonetic study of North American languages history and results. Ian Maddieson (Dept of Linguist., Univ. of California, Berkeley, CA 94720)

Serious phonetic study of North America languages started at the beginning of the 20th century. Within limits of available technology, aspects of speech articulation, aerodynamics and acoustics were investigated. One clear motivation was to understand how classes of sounds unfamiliar from study of better-known European and Asian languages were produced. Glottalized consonants and stops then referred to as "intermediate" (voiceless unaspirated) received particular attention. Nasal airflow, lip position in vowels, and tone and pitch accent were also investigated. Significant insights on relative timing were obtained, inter alia foreshadowing VOT measurement as a useful discriminator of laryngeal activity and revealing part of the mechanism by which ejective consonants are made. After the 1920's, the idea of "psychologically real" phonemes which ignored phonetic differences became the ruling paradigm in American linguistics, contributing to a decline of interest in phonetic studies that basically lasted until around the 1980's. When interest renewed, a new vision that phonetic patterns show regularities independent of phonemic structure guided research, and considerable attention was also paid to how indigenous American languages fit into overall phonetic typologies. Recent work is also often informed by concern for documentation of endangered languages and community interest in language revival.

8:30

3aSC2. Pitch accent timing in Chickasaw. Matthew K. Gordon (Dept. of Linguist., Univ. of California, Santa Barbara, Santa Barbara, CA 93106)

This paper examines the temporal realization of pitch accents in Chickasaw, a Muskogean language of Oklahoma. Questions in Chickasaw have a Lat their right edge preceded by a H* pitch accent that docks on one of the final three syllables of the last word in the Intonational Phrase: on a final CVV, otherwise on a heavy (CVV or CVC) penult, otherwise on the antepenult. Three speakers of Chickasaw were recorded uttering a list of words varying in the location of the pitch accent and the type of syllable carrying the pitch accent. Results indicate a tendency to realize the pitch accent relatively late in non-final pitch accented syllables (during the last 25 and during the coda consonant following a short vowel). However, final accented syllables realize the pitch accent early (within the first 20 vowel) and also display substantial vowel lengthening (by over 50 vowels in the penult or antepenult). These results suggest that pressure to avoid tonal crowding between the pitch accent and boundary tone overrides the late realization of pitch accents characteristic of positions in which crowding is less of an issue.

8:55

3aSC3. How the pharyngeal resonator is controlled articulatorily. John H. Esling (Dept of Linguist., Univ. of Victoria, Victoria, BC, Canada V8W 3P4, esling@uvic.ca)

Indigenous languages of North America, especially west coast Vancouver Island Nuuchahnulth and Interior BC Salish, illustrate an extensive range of use of the pharyngeal resonator. The pharynx is classified as part of the laryngeal articulator (rather than lingual) because the laryngeal constrictor mechanism (controlling changes from the glottis to the aryepiglottic folds) is the principal articulator whose movements determine the shape of the pharynx. The Nuuchahnulth (Nootka) dialects (e.g., Ahousaht) and Thompson, Moses-Columbia, and Spokane/Kalispel Salish varieties all close the larynx completely for the speech sound epiglottal stop. With the airway optimally shut, the pharynx is small; the aryepiglottic folds are pressed against the epiglottic tubercle, the tongue is retracted, and the larynx is raised. There are good indications that the thyroarytenoid, aryepiglottic, thyroepiglottic, hyoglossus and hyothyroid muscles participate in this buckling manoeuvre. These languages also contain either a pharyngeal fricative or approximant (at the same place of articulation but with less stricture) and, in Thompson, also pharyngealized uvulars (with even less stricture). Pharyngeal resonance in all of these cases is dependent initially on the shortening of the supraglottic tube and subsequently on the combined effect of tongue retraction and larynx raising reducing the volume immediately above the supraglottic tube.

9:20

3aSC4. One hundred years of instrumental phonetic fieldwork on North America Indian languages. Joyce McDonough (Dept. of Linguist., Univ. of Rochester, Rochester, NY 14627)

A resurgence of interest in phonetic fieldwork on generally morphologically complex North American Indian languages over the last 15 years is a continuation of a tradition started a century ago with the Earle Pliny Goddard, who collected kymographic and palatographic field-data between 1906–1927 on several Athabaskan languages: Coastal Athabaskan (Hupa and Kato), Apachean (Mescalero, Jicarilla, White Mountain, San Juan Carlos Apache), and several Athabaskan languages in Northern Canada (Cold Lake and Beaver); data that remains important for its record of segmental timing profiles and rare articulatory documentation in then largely monolingual communities. This data in combination with new work has resulted in the emergence of a body of knowledge of these typologically distinct families that often challenge notions of phonetic universality and typology. Using the Athabaskan languages as benchmark example and starting with Goddard's work, two types of emergent typological patterns will be discussed; the persistence of fine-grained timing and duration details across the widely dispersed family, and the broad variation in prosodic types that exists, both of which are unaccounted for by phonetic or phonological theories.

9:45

3aSC5. Coronal ejectives in Ahtna Athabaskan. Siri Tuttle (Univ. of Alaska, Fairbanks, AK 99775)

Ahtna, a non-tonal Alaskan language, is one of few Athabaskan languages that has retained word-final ejectives. The loss of stem-final glottal stops and ejectives in the other languages is correlated with tonogenesis. Hargus (2004) finds voice quality distinctions (creakiness, variability, and increased energy in higher harmonics) preceding stem-final glottal consonants in Witsuwiten, another nontonal Northern language. The present study considers the acoustics of stem-initial and stem-final [t'] and [ts'] in Central and Lower Ahtna. Stem-initial ejectives are canonically shaped, with defined silent period following oral release. In stem-final position, as predicted by Kari (1990), variable non-ejective realizations are found in both word-final and intervocalic (suffixed) position. These realizations are accompanied by variable voice quality effects in preceding vowels, more frequent and pronounced than those found variably in vowels following stem-initial ejectives. Stem-final [t'] is frequently pronounced as glottal stop, and final [ts'] as [s] where <, > stands for glottal stop. Strikingly, non-ejective realizations of stem-finals are found in suffixed and non-suffixed pairs with the same stem, suggesting that stem-final effects generalize to all tokens of stems. These findings support Leers (1979) picture of spirantization and suprasegmentalization in Athabaskan tonogenesis, as glottal effects gravitate to nucleus position.

10:10–10:25 Break

10:25

3aSC6. Tone and prosodic organization in Cherokee nouns. Keith Johnson (Dept. of Linguist., 1203 Dwinelle Hall, UC-Berkeley, Berkeley, CA 94720-2650) and Marcia Haag (Norman, OK 73019-2032)

Preliminary observations in the speech of one speaker of Cherokee led us to postulate three factors affecting tone in Cherokee. (1) Tone may be lexically specified with distinctive low, low fall, low rise, and high tones. (2) There is a metrically determined high fall pattern which may be distributed over not more than 2 syllables from the right edge of a prosodic domain. (3) Intonational domains may be associated with discourse functions, marked by high fall, or by pitch range upstep. This paper tests these observations in recordings of word lists and sentences produced by five additional speakers. The analysis we give, positing both lexical tone and metrical prosodic accent, is not unique in descriptions of language, but is different from the usual description of Cherokee. [Work supported by NSF.]

3a WED. AM

All posters will be on display and all authors will be at their posters from 10:50 a.m. to 12:00 noon.

3aSC7. Harmony in Gitksan. Noriko Yamane-Tanaka and Atsushi Fujimori (Dept. of Linguist., Univ. of University, E270-1866 Main Mall, Vancouver, BC, Canada V6T 1Z1)

What types of consonants are transparent in vowel harmony in Gitksan has been little understood. Previous studies suggest that it is translaryngeal harmony, where the qualities of unstressed suffixal vowels are harmonized with the ones of last stem vowels, apparently skipping the stem-final laryngeal stop and laryngeal fricative [Rigsby, unpublished manuscript, University of Queensland, Australia, 1986; Rigsby and Ingram, *International Journal of American Linguistics* 56, 251–263 (1990)]. However, our original data show that harmony takes place when the stem-final consonant is uvular as well. It suggests that it cannot simply be analyzed as laryngeal transparency, but it could be reanalyzed as [+low] transparency. The present study examines cross-consonantal vowel harmony through acoustic experiment. It is predicted that there is a significant difference in formant values between the vowels across uvular or laryngeal consonants versus the ones across the other consonants. New data from an Eastern Gitksan speaker (female, 72) and its analysis will be presented. Implications for locality of vowel harmony as well as for the phonological treatment of consonantal transparency will be discussed. [Work supported by SSHRC.]

3aSC8. The effect of [h] on tonal development in Nahuatl. Susan G. Guion (Linguist. Dept., 1290 U. of Oregon, Eugene, OR 97403, guion@uoregon.edu) and Jonathan D. Amith (Gettysburg College, Gettysburg, PA 17325)

In both its classical and modern forms, Nahuatl is usually described as having penultimate accent (with little discussion as to whether this is a phrase- or word-level phenomena). However, some varieties of Mexican Guerrero Nahuatl have recently been described as having word accent in locations related to the position of a historical non-word-final coda *h. Here, the hypothesis is investigated that coda *h lowered the fundamental frequency (F0) of the tautosyllabic vowel and that a H–L pitch contour was subsequently incorporated into the lexical representation of the word as *h was weakened and lost. Recordings were made of four dialects at varying stages of *h loss and tonal innovation. Six speakers from each dialect were recorded producing words in phrase medial and phrase final positions, some with historical coda *h and some without, and the F0 of each syllable in these words measured. The results indicated that the new accentual patterns found in some dialects have their origin in the phonetic lowering of F0 in syllables containing coda *h. It also seems that once a new accentual pattern is developed and the *h lost, the resulting phonologized pattern may not correspond to the phonetic locus of the sound change.

3aSC9. Glottalization cues of Nuu-chah-nulth glottalized resonants. Marion G. Caldecott (Dept. of Linguist, Univ. of British Columbia, Canada V6T 11, mariongc@interchange.ubc.ca)

Theories of positional neutralization (Steriade, 1997) predict that pre-glottalized resonants should not occur word-initially because the cues to glottalization are not salient. Contrary to this, Nuu-chah-nulth, a Wakashan language, has pre-glottalized that do occur in word-initial position. This paper will show that a severely truncated resonant portion is the cue to glottalization in resonants in Nuu-chah-nulth. Previous research [Carlson *et al.*, *J. Jpn. Int. Prop. Assoc.* 31, 275–279 (2001)] shows that Nuu-chah-nulth glottalized resonants consist of a stop closure that is phonetically identical to phonemic glottal stop followed by a (mostly) plain resonant. That study also found that glottalized resonants are twice as long in initial position as plain resonants. Following up on this, a pilot study [Caldecott and Kim, presented at the 1st annual Wakashan Linguistics Conference,

2004] indicated that the length of the resonant portion in glottalized resonants was in fact half that of a normal resonant. This paper expands on the previous study, incorporating new acoustic data on resonants in all word positions. Data are elicited from a fluent native speaker and measured using Praat. [Work supported by SSHRC.]

3aSC10. Glottalized resonants in NLe kepmxcin (Thompson River Salish). Patricia A. Shaw, Fiona M. Campbell (Dept. of Linguist., Univ. of British Columbia, E270–1866 Main Mall, Vancouver, BC, Canada V6T 1Z1), and Flora Ehrhardt (Lytton Indian Band, BC)

This study examines properties of glottalized resonants (/m', n', l', y', w'/) in NLe?kepmxcin, a Northern Interior Salish language, both comparing them with their plain counterparts (m, n, l, y, w) and considering contextual factors as defined by word, syllable, and prominence positions. Despite the relative cross-linguistic rarity of glottalized resonants, they are of considerable theoretical interest given that several recent studies [e.g., Bird and Caldecott, *Proceedings of the 10th AICSST* (in press)] indicate that their articulatory and acoustic realization may vary significantly across languages and contexts. Of particular relevance to the present study is Carlson, Esling, and Harris' [UMOPL 17, 58–71 (2004)] claim that in NLe?kepmxcin “glottalized resonants are a sequential combination of the resonant plus [glottal stop]” [pg. 64]. Based on the analysis of glottalized resonants in this study, it appears that the picture is more complex than consistent post-glottalization. Of special interest are glottalized resonants in word-initial position, both because this class is seriously under-represented in the previous literature, and because its characteristics challenge Carlson, Esling, and Harris' sequencing generalization. Understanding the organization of glottalized resonants will help clarify both sequential and markedness issues in NLe?kepmxcin phonology. [Work supported by SSHRC.]

3aSC11. Phonetic structures of Washo. Patrick J. Midtlyng and Alan C. L. Yu (Dept. of Linguist., Univ. of Chicago, 1010 E 59 St., Chicago, IL 60637, aclyu@uchicago.edu)

Washo is a severely moribund language spoken in several townships near the California-Nevada border southeast of Lake Tahoe. The Washo language is now used only by approximately 13–20 elderly speakers with very few fluent speakers under 60. Recordings were collected in Dressler-ville, Nevada from two native speakers. Recordings collected by William Jacobsen Jr. in the 1960s provided additional data for comparison. Analysis showed that all the stops were either voiceless unaspirated, voiceless aspirated, or ejectives. Vowel analyses showed the nature of the six long and short vowels. Special attention was paid to the nature of vowels and consonants in and around the stressed syllable. Post-tonic sibilants and sonorants are lengthened intervocalically even though consonantal length is not contrastive in Washo. Vowel length contrasts only in the stressed syllable. Stressed vowels are always long before an intervocalic voiceless unaspirated stop, but vocalic length contrast remains before aspirated stops and ejectives.

3aSC12. An ultrasonic investigation of retracted consonants in St'at'imcets (Lillooet Salish). Nahal Namdaran (Dept. of Linguist., Univ. of British Columbia, 1866 Main Mall, Buchanan E270, Vancouver, BC, Canada V6T 1Z1)

Tongue retraction is rare in native languages of North America. Previous literature on St'at'imcets (also known as Lillooet) has left unanswered the question of how retracted consonants (uvulars, a pharyngeal, and a subset of retracted coronals) are articulated, and whether they in fact involve tongue retraction. Based on perceptual and acoustic data, previous descriptions differ with respect to the involvement of the tongue dorsum

and tongue root in the articulation of St'at'imcets retracted consonants, particularly with the orthographic “g,” disputed to be either a uvular or pharyngeal approximant [Shahin, 1997; van Eijk, 1997]. Articulatory data of retracted and nonretracted consonants were collected from 4 St'at'imcets speakers via a portable ultrasound machine. Results indicate a significant difference in tongue root, dorsum, and body position for retracted versus nonretracted consonants, with the former possessing more tongue root and dorsum backing and tongue body lowering than the latter. Furthermore, results show “g” to be articulated similarly to the uvular consonant, with the tongue dorsum moving towards the posterior uvula/upper pharyngeal region of the vocal tract. These results provide direct evidence for the articulatory makeup of St'at'imcets retracted consonants and shed light on St'at'imcets retraction effects and phonological harmony processes. [Work supported by SSHRC.]

3aSC13. An acoustic study of Deg Xinag fricatives. Richard Wright, Sharon Hargus, and Julia Miller (Dept. of Linguist., Univ. of Washington, Box 354340, Seattle, WA 98195-4340)

Few studies have looked at the acoustic properties of fricatives in Native American languages. Notable exceptions are McDonough (2003) on Navajo, and Gordon, Barthmaier, and Sands (2002), which examined a variety of languages, including several Native American ones. In Deg Xinag, an endangered Athabaskan language spoken in Alaska, there are eight places of articulation for voiceless fricatives [a relatively large phonetic inventory, according to Maddieson (1984)], including some rarely studied place contrasts (e.g. alveolar versus retroflex). In this study, pre- and post-vocalic fricatives were digitally recorded in the field from eight speakers (two males, six females) using a head-mounted mic to control for distance from the source. The segmental context was also controlled for, the neighboring vowel being [a] in all cases. Each speaker produced four repetitions of each word. The first four spectral moments, lowest spectral peak, and the three loudest spectral peaks were measured at the midpoint of each fricative, and each speaker's average for each of the measures was calculated. In this poster, qualitative results in the form of spectrographic analysis will be presented. Repeated measures ANOVA for each of the quantitative measures will also be presented. [Work supported by NSF.]

3aSC14. Glottalized consonants in Oaxaca Chontal. Heriberto Avelino and Ian Maddieson (Dept. of Linguist., Univ. of California, Berkeley, CA 94720)

Lowland Oaxaca Chontal, a language generally classified as Hokan and hence related to Yuman and Pomo languages, is now spoken by only a small percentage of the Chontal people who live on the Pacific coast of Oaxaca state, Mexico. Recordings were made of seven speakers in their 60's to 80's in December 2003. The language has a series of glottalized obstruents some of which vary in production between fricative and affricate realizations. Our data show the labio-dental and lateral are the most likely to have fricative realizations, the coronals are regularly affricates and the velar is regularly a stop. These segments are contrastive but morphologically related to the voiceless fricatives of the language. This is demonstrated by one common plural formation process which adds glottalization to a final consonant (e.g. /apix/ “stone,” /apik'/ “stones”). After a vowel the same morpheme is realized as a glottal stop (e.g. /u/ “eye,” /u/ “eyes”). However, not all sequences of voiceless fricative + glottal stop are realized as glottalized obstruents (e.g. /ase/ “atole”). Chontal also has glottalized nasals and approximants. As in other languages the accompanying glottal constriction is variable in degree and timing.

3aSC15. Consonantal perturbations of pitch in Halkomelem Salish. Jason Brown and James J. Thompson (Dept. of Linguist., Univ. of British Columbia, 1866 Main Mall, Buchanan E270, Vancouver, BC, Canada V6T 1Z1, jcb@interchange.ubc.ca)

It has long been noted that consonants have an effect on the pitch of a following vowel: voiceless stops tend to raise F_0 , while voiced stops lower it. It has also been suggested that the duration of such perturbations is shorter in tone languages than in non-tone languages [Hombert, *Studies in African Linguistics* (1977)]. This study compares the effects that consonants have on F_0 in two closely related Salish languages: Island Halkomelem, a non-tone language, and Upriver Halkomelem, a language that has reportedly undergone some limited tonogenesis but offers no clear prosodic clues regarding tonality. The effects of the voiceless and ejective stop series were observed, and measurements of F_0 were taken at the onset of voicing for the vowel, then at 20 msec. intervals up to 100 msec. Preliminary results indicate that i) Island Halkomelem shows a greater magnitude of difference in F_0 at vowel onset between the voiceless and ejective stops than Upriver Halkomelem, and ii) Island Halkomelem shows greater durations of consonantal perturbations of F_0 than does Upriver Halkomelem. This suggests that Upriver Halkomelem may have become more sensitive to pitch than the Island dialect, supporting the interpretation of this language as tonal. [Work supported by Phillips Fund.]

3aSC16. Acoustic cues to Nehiyawewin constituency. Clare Cook and Jeff Muehlbauer (Univ. of British Columbia, Dept. of Linguist., 1866 Main Mall, Buchanan E270, Vancouver, BC, Canada V6T 1Z1, clarec@interchange.ubc.ca)

This study examines how speakers use acoustic cues, e.g., pitch and pausing, to establish syntactic and semantic constituents in Nehiyawewin, an Algonquian language. Two Nehiyawewin speakers autobiographies, which have been recorded, transcribed, and translated by H. C. Wolfart in collaboration with a native speaker of Nehiyawewin, provide natural-speech data for the study. Since it is difficult for a non-native-speaker to reliably distinguish Nehiyawewin constituents, an intermediary is needed. The transcription provides this intermediary through punctuation marks (commas, semi-colons, em-dashes, periods), which have been shown to consistently mark constituency structure [Nunberg, CSLI 1990]. The acoustic cues are thus mapped onto the punctuated constituents, and then similar constituents are compared to see what acoustic cues they share. Preliminarily, the clearest acoustic signal to a constituent boundary is a pitch drop preceding the boundary and/or a pitch reset on the syllable following the boundary. Further, constituent boundaries marked by a period consistently end on a low pitch, are followed by a pitch reset of 30–90 Hz and have an average pause of 1.9 seconds. I also discuss cross-speaker cues, and prosodic cues that do not correlate to punctuation, with implications for the transcriptional view of orthography [Marckwardt, Oxford 1942].

3aSC17. Discourse-level contours in Nehiyawewin. Jeff Muehlbauer and Clare Cook (Univ. of British Columbia, Dept. of Linguist., 1866 Main Mall, Buchanan E270, Vancouver, BC, Canada V6T 1Z1, jefmuehl@interchange.ubc.ca)

This study describes declination and discourse-sized intonation contours in Nehiyawewin, an Algonquian language whose pitch and intonation systems have not been previously studied. The study draws on 270 min of recordings of two female Nehiyawewin elders telling their life stories to another Nehiyawewin native speaker. Data is analyzed by using Praat's default algorithm to generate f_0 curves for each breath group. Preliminary results indicate: (1) When breath-group internal pitch peaks are considered, an obvious downward trend of f_0 occurred in fewer than half the breath groups analyzed (about 40% or 37/90). This raises questions about the role of classical declination in natural discourse [Umeda, *Journal of*

Phonetics **10** (1982)]. (2) When we abstract away from declination within a breath group by computing mean f_0 and mean pitch peak for each breath group and tracking trends for these means, larger patterns seem to emerge; breath groups can be grouped into larger units based on raising

and lowering trends. These units have a mean peak range of 150 Hz with a 30 Hz change from one breath group to the next and correspond to a domain of around five clauses (about 3–4 breath groups, about 45 syllables).

WEDNESDAY MORNING, 18 MAY 2005

GEORGIA A, 9:00 A.M. TO 12:00 NOON

Session 3aSP

Signal Processing in Acoustics: Potpourri

David M. Fromm, Chair

Naval Research Laboratory, 4555 Overlook Ave., SW, Washington, DC 20375-5350

Contributed Papers

9:00

3aSP1. Wavelet based detection of manatee vocalizations. Berke M. Gur and Christopher Niezrecki (Dept. of Mech. Eng., Univ. of Massachusetts Lowell, One University Ave., Lowell, MA 01854-2881, Chris_Niezrecki@uml.edu)

The West Indian manatee (*Trichechus manatus latirostris*) has become endangered partly because of watercraft collisions in Florida's coastal waterways. Several boater warning systems, based upon manatee vocalizations, have been proposed to reduce the number of collisions. Three detection methods based on the Fourier transform (threshold, harmonic content and autocorrelation methods) were previously suggested and tested. In the last decade, the wavelet transform has emerged as an alternative to the Fourier transform and has been successfully applied in various fields of science and engineering including the acoustic detection of dolphin vocalizations. As of yet, no prior research has been conducted in analyzing manatee vocalizations using the wavelet transform. Within this study, the wavelet transform is used as an alternative to the Fourier transform in detecting manatee vocalizations. The wavelet coefficients are analyzed and tested against a specified criterion to determine the existence of a manatee call. The performance of the method presented is tested on the same data previously used in the prior studies, and the results are compared. Preliminary results indicate that using the wavelet transform as a signal processing technique to detect manatee vocalizations shows great promise.

9:15

3aSP2. Background noise cancellation for improved acoustic detection of manatee vocalizations. Zheng Yan (Dept. of Mech. and Aersp. Eng., Univ. of Florida, Gainesville, FL 32611-6250), Christopher Niezrecki (Univ. of Massachusetts Lowell, Lowell, MA 01854-2881, Chris_Niezrecki@uml.edu), and Diedrich O. Beusse (Univ. of Florida, Gainesville, FL 32610-0126)

The West Indian manatee (*Trichechus manatus latirostris*) has become endangered partly because of an increase in the number of collisions with boats. A device to alert boaters of the presence of manatees, so that a collision can be avoided, is desired. Practical implementation of the technology is dependent on the hydrophone spacing and range of detection. These parameters are primarily dependent on the manatee vocalization strength, the decay of the signal strength with distance, and the background noise levels. An efficient method to extend the detection range by using background noise cancellation is proposed in this paper. An adaptive line enhancer (ALE) that can detect and track narrowband signals buried in broadband noise is implemented to cancel the background noise. The results indicate that the ALE algorithm can efficiently extract the manatee calls from the background noise. The improved signal-to-noise ratio of the

signal can be used to extend the range of detection of manatee vocalizations and reduce the false alarm and missing detection rate in their natural habitat.

9:30

3aSP3. Electronic noise suppression in the wavelet, frequency, and time domains. Juan Arvelo, Jr. (Appl. Phys. Lab., Johns Hopkins Univ., 11100 Johns Hopkins Rd., Laurel, MD 20723-6099)

Automatic machine gun recordings were severely contaminated with electronic noise. A detailed examination of the character of the electronic noise revealed that it consists of periodic pulses at an interval of about 4.2 ms and that its frequency content is rich in narrow-band lines with a colored broadband background. Given the impulsive nature of the electronic noise in the time domain, wavelet decomposition was applied in an attempt to harness the noise and to null it out. An alternative approach was also implemented where a bank of stop band Butterworth filters were applied in the frequency domain to suppress the narrow-band components of the electronic noise. Other approaches include the use of adaptive cancellation in the time and frequency domains. We will demonstrate and compare the performance of these approaches emphasizing the need to minimize their influence in distorting the signal of interest.

9:45

3aSP4. Maximum entropy method for constructing environmental parameter probability density functions. Leon H. Sibul, R. Lee Culver, David L. Bradley, and H. John Camin (Appl. Res. Lab. and Grad. Program in Acoust., P.O. Box 30, State College, PA 16804, rlc5@psu.edu)

At-sea experience has shown that target detection ranges are often less than predicted using current sonar system performance prediction methods. One possible reason for this disparity is that classical, sonar equation-based performance prediction models do not take into account the random and uncertain nature of signal propagation and scattering in a real ocean environment. Our goal is to develop a probabilistic sonar performance prediction methodology that can make use of limited knowledge of random or uncertain environment, target and sonar system parameters, but does not make unwarranted assumptions or rely upon unknown information. Beginning with models that describe how the ocean environment effects the received signal, the Maximum Entropy method (MEM) can be used to construct probability density functions (PDFs) for the parameters of those models. The benefit of the MEM is that it allows one to incorporate available knowledge about the parameters, but is maximally non-committal about what is not known about the parameters. The methodology is presented and illustrated using a model for particular environmental effects on the structure of low-frequency signals propagating over long distances in the ocean. [Work supported by ONR Code 321US.]

10:15

3aSP5. Stochastic filtering. David I. Havelock (Natl. Res. Council, 1200 Montreal Rd., Ottawa, ON, Canada K1A 0R6, david.havelock@nrc.ca)

Digital filters convolve a signal with an impulse response. This is equivalent to computing a weighted sum of delayed versions of the signal, with the signal versions indexed by their sample-delay and the weights given by the filter impulse response. This talk explores a different approach to filtering. Rather than computing a weighted sum of data values, the *stochastic filter* uses a sequence of random indices to select output data values from delayed versions of the signal. No summation is done. The probability distribution of the random indices is made identical to the normalized modulus of the impulse response of the corresponding conventional filter. For indices with negative weights, data values are negated. For a random index sequence $s(n)$ the stochastic filter output is simply $y(n) = \pm x(n - s(n))$. The output of such a stochastic filter is noisy but it is *equal in the mean* to the output of the corresponding conventional filter. If the sequence of random indices is identically and independently distributed then the noise is spectrally white. Such a filter can be useful for estimating spectral parameters with minimal computation. The principle is illustrated by an application to double-talk detection for telephones.

10:30

3aSP6. Autonomous signal separation microphone system based on the spatio-temporal gradient analysis. Kenbu Teramoto and Kosuke Tsuruta (Dept. of Mech. Eng., SAGA Univ., SAGA, 8408502, Japan)

This paper presents a novel autonomous directivity microphone system based on the newly proposed spatio-temporal blind source separation. Recently, the methods of blind source separation and multi-deconvolution of source signals have been proposed in various fields, especially, acoustic applications including the cocktail-party problem. The blind source separation principally uses no *a priori* knowledge about parameters of convolution, filtering and mixing. In the simplest case of the blind source separation problems, observed mixed signals are linear combinations of unknown mutually statistically independent, zero-mean source signals. The spatio-temporal blind signal separation algorithm utilizes the linearity among the four signals: (1) temporal gradient, (2) x ; (3) y ; and (4) z -directional spatial gradients of the sound-pressure, all of which are governed by the equation of advection. The proposed method, therefore, has an ability to simplify the convolution blind source separation problems into the instantaneous blind source separation over the spatio-temporal gradient space. The acoustical experiments are performed with the particle velocity microphone (Microflown) successfully instead of sound pressure microphones. Because, x ; y ; and z -directional gradients of the sound pressure are equivalent the temporal derivatives of the corresponding directional particle velocities theoretically.

10:45

3aSP7. Bayesian sampling with applications to energy decay analysis in acoustically coupled spaces. Tomislav Jaska and Ning Xiang (School of Architecture and Dept. of Electrical, Computer, and Systems Eng., Rensselaer Polytechnic Inst., Troy, NY 12180)

Bayesian inference comprises of parameter estimation and model selection/comparison. A common approach to solving both of these problems has been to use statistical sampling techniques that are inherently non-Bayesian. This paper presents new Bayesian sampling method that solves both problems by changing the focus of Bayesian inference to the model selection problem first. In the papers [N. Xiang and P. M. Goggins, *J. Acoust. Soc. Am.* **110**, 1415–1424 (2001); **113**, 2685–2697 (2003)], the authors developed a model for the decay times and decay modes of acoustically coupled rooms in terms of measured Schroeder's decay functions. This paper shows how the Bayesian sampling method can be used to evaluate the "Bayesian evidence" term used in model selection as well as determining the decay times along with error estimates.

3aSP8. Adaptive Schur algorithm dedicated to underwater transient signal processing. Maciej Lopatka (Lab. d'Informatique Industrielle et Automatique, Univ. Paris XII, 61 av Gaille, 94010 Créteil, France), Olivier Adam, Christophe Laplanche (Univ. Paris XII, France), Jan Zarzycki (Tech. Univ. of Wroclaw, Poland), and Jean-François Motsch (Univ. Paris XII, France)

The algorithm proposed by Lee and Morf [*IEEE Transactions on Circuits and Systems* **28**(6) (1981)] which stems from the method defined by Schur [*Operator Theory: Advances & Application*, Vol. 18 (Birk-Verlag, 1986)] has acquired a new significance [Zarzycki, *Journal of Multidimensional Systems & Signal Processing* (Kluwer Academic, 2004)], due to its performances and particularly due to its applications in real time, made possible by the speed of processors available nowadays. Based on the innovations filter principle, Schur's proposal models the signal by calculating reflection coefficients, describing entirely the second-order signal. The reflection coefficients can be simply transformed to the AR coefficients, from which one derives the time-frequency representation. We compare performances of this approach with other time-frequency representations commonly used in the signal processing (spectrogram, AR, wavelet transform); we subsequently present the results obtained for transitory underwater acoustic signals, which our laboratory is investigating. The Lee and Morf algorithm offers an excellent tracking of the signal's characteristics and allows us to systematically detect transitory signals. This is particularly pertinent to segmentation problems relating to the application of underwater acoustics. The robustness of the Schur detector and a resolution of the Schur time-frequency representation support the resurgence of the Schur algorithm.

11:15

3aSP9. Laser Doppler signal for high amplitude sound wave. Jean-Christophe Valiere, Solenn Moreau, Helene Bailliet (LEA-CNRS UMR 6609, Université de Poitiers, France), Philippe Herzog (LMA-CNRS, Marseille, France), and Laurent Simon (LAUM-CNRS 6613, Université du Maine, France)

Laser Doppler velocimetry (LDV) is now recognized as a very useful technique for estimating acoustic velocity with a good time resolution in many applications. Previous research showed that the most important parameter in LDV for acoustics is the particle displacement and particularly its magnitude compared with the size of the probe volume formed by the crossing of the laser beams. Specific techniques were developed to estimate the acoustic wave when the displacement is of the same order of magnitude as the probe diameter and when it is much smaller. In this study, we investigate situations where the displacement is much higher than the probe volume. The measuring process has been simulated numerically and it appears that the process leads to an under-estimation of the velocity around zero. This under-estimation is due to the processing of the laser Doppler signal in the case of high displacements. The associated signal is a non poissonnian randomly sampled signal to which classical processing methods are not adapted. It is compared to experimental signals obtained in the context of a study of non linear effects in an acoustic wave guide. To this end, different processing of both simulated and experimental signals are presented and compared.

11:30

3aSP10. Cross-frequency Doppler sensitive signal processing. Ronald A. Wagstaff (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, Univ., MS 38677, rwagstaf@olemiss.edu)

When there is relative motion between an acoustic source and a receiver, a signal can be Doppler shifted in frequency and enter or leave the processing bins of the conventional signal processor. The amount of the shift is determined by the frequency and the rate of change in the distance between the source and the receiver. This frequency Doppler shifting can cause severe reductions in the processors performance. Special cross-frequency signal processing algorithms have recently been developed to mitigate the effects of Doppler. They do this by using calculation paths that cut across frequency bins in order to follow signals during frequency

shifting. Cross-frequency spectral grams of a fast-flying sound source were compared to conventional grams, to evaluate the performance of this new signal processing method. The Doppler shifts in the data ranged up to 70 contiguous frequency bins. The resulting cross-frequency grams showed that three paths provided small to no improvement. Four paths showed improvements for either up-frequency or down-frequency shifting, but not for both. Two paths showed substantial improvement for both up-frequency and down-frequency shifting. The cross-frequency paths will be defined, and comparisons between conventional and cross-frequency grams will be presented. [Work supported by Miltec Corporation.]

11:45

3aSP11. Eliminating clutter by coordinate zeroing. Harry DeFerrari (RSMAS—Univ. of Miami, 4900 Rickenbacker Cswy, Miami, FL 33149) and Andrew Rodgers (Woods Hole Group, East Falmouth, MA 02536)

Shallow water doppler sonar is limited by clutter that masks the presence of slow targets. Clutter arises from scattering from irregularities in

the bottom or from fish schools. If the source and receiver are fixed in a bi-static sonar configuration, then the bottom scatterers (and sometimes the fish) have only a zero doppler component. Clutter arises from the doppler leakage resulting from sonar signals that are inherently short in time and therefore cannot resolve bottom reverberation in frequency. Here long continuous pulse compressions signals are considered (m-sequences). Ambiguity diagrams are compared with simple CW pulse and PCM pulses. The unique correlation properties of M-sequences lead to a signal sampling method that resolves sonar returns in complete ortho-normal (CON) data sets. Zero doppler returns are identified as either direct arrivals or reverberation returns and then their particular coordinate can be set zero (coordinate zeroing) without affecting other data points. In this way, the zero-doppler bottom reverberation and all doppler leakage is eliminated. As an example, the return from a slow doppler target is imbedded in a very large number of reverberation returns that have higher signal level. As the zero doppler contributions are removed by coordinate zeroing the target becomes detectable.

WEDNESDAY MORNING, 18 MAY 2005

GEORGIA B, 8:15 A.M. TO 12:00 NOON

Session 3aUW

Underwater Acoustics and Acoustical Oceanography: Frequency Dependence of Sound Speed and Attenuation in Marine Sediments I

Gopu Potty, Cochair

Dept. of Ocean Engineering, Univ. of Rhode Island, Narragansett, RI 02882

Peter H. Dahl, Cochair

Applied Physics Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698

Chair's Introduction—8:15

Invited Papers

8:20

3aUW1. Nonlinear frequency dependence of the effective seabottom acoustic attenuation from low-frequency field measurements in shallow water. Ji-Xun Zhou and Xue-Zhen Zhang (Georgia Inst. of Technol., Atlanta, GA 30332-0405 and Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC)

Hamilton's seabed geo-acoustic model, which is widely accepted, predicts that the attenuation of sound in marine sediments increases linearly with frequency over the full frequency range of interest in ocean acoustics (a few hertz to megahertz). However, Biot-Stoll's physics-based seabed geo-acoustic model predicts that the bottom attenuation should exhibit non-linear frequency dependence, particularly in sandy bottoms. Since the publication of previous papers [Zhou, *J. Acoust. Soc. Am.* **78**, 1003–1009 (1985); Kibblewhite, *ibid.* **86**, 716–738 (1989)], more low-frequency field data, collected from different coastal zones around the world, have shown inconsistencies with the often-used linear frequency dependence. This paper attempts to support the non-linear frequency dependence of sea bottom attenuation through a review of shallow-water acoustic field measurements, with a special emphasis on the 50–1000 Hz frequency range. The relevant measurements include bottom reflection loss, sound transmission loss, dispersion analysis, vertical coherence of both propagation and reverberation, normal-mode spatial filtering, and optimum frequency and transition range of sound propagation. A non-linear frequency dependence of equivalent bottom acoustic attenuation, derived from these measurements, will be introduced. [Work supported by ONR and NNSF of China.]

8:40

3aUW2. Overview of SAX99 and SAX04 measurements of sediment sound speed and attenuation. Eric I. Thorsos, Kevin L. Williams, Darrell R. Jackson, and Dajun Tang (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, eit@apl.washington.edu)

During SAX99 (for sediment acoustics experiment—1999) the sediment sound speed (125 Hz to 400 kHz) and attenuation (2.5 to 400 kHz) in sandy sediments were measured by a variety of techniques. The SAX99 site was 2 km from shore on the Florida Panhandle near Fort Walton Beach in water of 18–19 m depth. SAX04 was held in the fall of 2004 at a site close to the SAX99 site, about 1 km from shore in water of 17 m depth. The sediment sound speed and attenuation were again measured over a broad

frequency range by multiple techniques, with even more attention paid to the low frequency band from 1–10 kHz. The results and corresponding uncertainties from SAX99 will be reviewed, and the consistency with Biot model predictions and alternative models (e.g., Buckingham's model) will be discussed. An overview will then be presented of the recently completed SAX04 measurement program on sediment sound speed and attenuation. [Work supported by ONR.]

9:00

3aUW3. Models for attenuation in marine sediments that incorporate structural relaxation processes. Allan D. Pierce, William M. Carey (Boston Univ., Boston, MA 02215, adp@bu.edu), and James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Biot's model leads to an attenuation coefficient at low frequencies that is proportional to ω^2 , and such is consistent with physical models of viscous attenuation of fluid flows through narrow constrictions driven by pressure differences between larger fluid pockets within the granular configuration. Much data suggests, however, that the attenuation coefficient is linear in ω for some sediments and over a wide range of frequencies. A common model that predicts such a dependence stems from theoretical work by Stoll and Bryan [J. Acoust. Soc. Am. **47**, 1440 (1970)], in which the elastic constants of the solid frame are taken to be complex numbers, with small constant imaginary parts. Such invariably leads to a linear ω dependence at sufficiently low frequencies and this conflicts with common intuitive notions. The present paper incorporates structural relaxation, with a generalization of the formulations of Hall [Phys. Rev. **73**, 775 (1948)] and Nachman, Smith, and Waag [J. Acoust. Soc. Am. **88**, 1584 (1990)]. The mathematical form and plausibility of such is established, and it is shown that the dependence is as ω^2 at low frequencies, and that a likely realization is one where the dependence is linear in ω at intermediate frequency ranges.

9:20

3aUW4. Propagation and reverberation implications of seabed dispersion. Henrik Schmidt (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

The significance of the seabed geoacoustic properties to shallow water acoustic propagation is well established, and consequently the dispersion, or frequency dependence of the sound speed and the attenuation may be significant to the performance of broadband acoustic systems, whether for acoustic sensing or for communication. On the other hand, it is equally well established that the seabed significance is highly dependent on the system configuration, e.g., source and receiver depths, and the same will therefore obviously be the case for the dispersion. This paper will review the propagation effects of frequency dependent seabed geoacoustics, under characteristic seasonal environmental conditions, and typical system configurations. The associated implications to the acoustic system performance will be discussed in the context of theoretical performance bounds, such as the Cramer-Rao lower bounds, using established wave theory propagation models for range-independent and -dependent environments. Among the specific examples will be the effect of strong sediment dispersion in sandy sediments on buried target detection by low-frequency SAS systems operating in the 1–10 kHz regime, and broadband acoustic communication systems, including the effect of dispersion on seabed reverberation. [Work supported by ONR.]

Contributed Papers

9:40

3aUW5. Attenuation in the sediment layers in East China Sea. Gopu R. Potty and James H. Miller (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett, RI 02882, potty@oce.uri.edu)

Compressional wave attenuation in the frequency range 50 to 350 Hz is estimated using broadband data collected in the East China Sea. A time-frequency analysis provides amplitudes of individual modes which are used to estimate the modal attenuation coefficients. Different modes are sensitive to sediment properties at different depths and this mode-depth sensitivity is explored in a sensitivity study. Based on this sensitivity study attenuation in different layers of the sediments will be estimated using different modes which are sensitive at these depths. This is a variation of previous studies where different frequencies have been used to invert for different sediment parameters using sub-space approaches. This approach focuses the inversion to various layers (provided by the geophysical surveys at the site) of the sediment and hence presents a tool to describe the depth variations in the sediment. The attenuation versus frequency behavior of the different sediment layers will be discussed. [Work supported by ONR.]

9:55–10:15 Break

10:15

3aUW6. Measuring dispersion and attenuation in sediment at low frequency. Steven A. Stotts, Robert A. Koch, and David P. Knobles (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029, stotts@arlut.utexas.edu)

Much work has been done at frequencies above 1 kHz to measure sediment dispersion and attenuation and these results have been analyzed within the context of various models such as Biot-Stoll, Buckingham, and BICSQS. Some effort has been made to establish the frequency dependence of sediment attenuation below 1 kHz, but few results are available concerning dispersion in that frequency band. The present work examines a technique for measuring low frequency sediment dispersion and attenuation with a real dataset. Recently, data were collected off the coast of Florida using implosive broadband sources to perform geo-acoustic inversions. Previous inversion studies with this data ignored any dispersion effects. [Stotts *et al.*, J. Acoust. Soc. Am. **115**, 1078–1102 (2004).] The environment was shown to consist of thin, hard, porous material less than 1 m thick overlying sand sediment. A ray trace plus plane-wave reflection coefficient model was used for the analysis. Here, dispersion is repre-

sented phenomenologically over a finite frequency band in the model by allowing both the sound speed and attenuation to be frequency dependent in both layers of a two-layer sediment plus substrate model.

10:30

3aUW7. Frequency dependence of posterior probability distributions for Biot parameters. Tracianne B. Neilsen, Marcia Isakson (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029, neilsen@arlut.utexas.edu), and Buye Xu (Brigham Young Univ., Provo, UT 84602)

The Biot-Stoll poro-elastic model for sediments is parameterized by 13 properties. Initial investigations indicate that the individual Biot parameters influence reflection loss differently depending on the frequency range of interest. A more extensive study of the frequency dependence of the parameter sensitivities is performed by computing posterior probability distributions (PPDs) [S. E. Dosso, *J. Acoust. Soc. Am.* **111**, 129–142 (2002)] to determine which parameters most influence reflection loss for each frequency decade from 100 Hz to 1 MHz. The sampling for the PPDs is based on the least-squares error between simulated reflection coefficient data and values modeled with the reflection module of OASES. [H. Schmidt, "OASES Version 2.1 User Guide and Reference Manual," Department of Ocean Engineering, Massachusetts Institute of Technology, 1997.] The PPDs provide a quantitative value of the uncertainty in the optimum parameter estimates obtained by sampling or inversion techniques. A good value of the uncertainty is critical if one wants to use parameter estimates obtained with data from one frequency range to predict acoustic response for a different frequency range. [Work supported by the ONR.]

10:45

3aUW8. A hybrid model of sound propagation in unconsolidated sediments. Brian T. Hefner and Kevin L. Williams (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

Efforts to model sound speed and attenuation in sandy sediments have centered on the use of theories for which either the relative motion of the pore fluid is the dominant attenuation mechanism, such as Biot theory, or the dominant loss mechanism is grain-to-grain friction. A recent model which attempts to incorporate grain-to-grain loss mechanisms into a model of sandy sediments was proposed by Buckingham. This model can fit the frequency dependence of the attenuation measured in ocean sediments and laboratory glass bead sediments, but it does not capture the sound speed dispersion as effectively as Biot theory. The relative success of each model suggests that both attenuation mechanisms may play important roles in sediment acoustics. In order to explore this possibility, a hybrid model has been developed which incorporates Buckingham's grain-to-grain shearing mechanisms into the frame moduli used in Biot theory. In the hybrid model, the grain-to-grain losses dominate at high and very low frequencies while pore fluid attenuation dominates at mid-frequencies where the sound speed dispersion is the most pronounced. As a consequence, the hybrid model is able to describe both the measured sound speed and attenuation in ocean and laboratory sediments. [Work supported by ONR.]

11:00

3aUW9. Attenuation in sand sediments due to scattering from force chains. Brian T. Hefner and Darrell R. Jackson (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

The stresses in a granular media, such as an ocean sediment, are known to be heterogeneous and tend to form chains along which the stresses are particularly high. Because the elastic properties of a granular medium depend nonlinearly on the static forces between the grains, these force chains should induce inhomogeneities in the frame moduli. The correlation lengths of these inhomogeneities have been observed to be on the order of only a few grain diameters and as a consequence have not been thought to produce any significant scattering of acoustic energy. However, for high frequency sound in a saturated granular medium, these inhomogeneities should be on the order of the slow compressional wave. As a

consequence, as the fast wave propagates through the sediment, it may scatter energy into the slow wave which is then rapidly attenuated. This scattering process may account for the deviation of the measured attenuation from the predictions of Biot theory that have been observed in ocean and laboratory sediments at high frequencies. We present a model of sound propagation through sand sediments that incorporates this loss mechanism and compare the results of this model to recent measurements. [Work supported by ONR.]

11:15

3aUW10. Sediment layer perturbation influences on the frequency dependence of effective bottom attenuation. James D. Nickila (Adv. Sonar Technol. Div. Naval Undersea Warfare Ctr., Newport, RI 02841), Kevin B. Smith (Naval Postgrad. School, Monterey, CA 93943), and Gopu Potty (Univ. of Rhode Island, Narragansett, RI 02882)

Over the past several years, concern has grown over the appropriateness of bottom attenuation models that assume a linear frequency dependence. Empirical analyses of experimental data have suggested power law dependence with frequency exponents as high as 1.7 and above, but with large variability between geographic regions. [Zhou *et al.*, *J. Acoust. Soc. Am.* **82**, 287–292 (1987).] The fundamental cause of this dependence is unknown. In this study, we investigate the effect of perturbations (both random and deterministic) in the sediment layer on the effective attenuation factor for a range of frequencies. A parabolic equation (PE) model is used to generate the acoustic field for a given environmental realization. The pressure field is decomposed using a normal mode expansion and an effective attenuation factor is calculated. Regression is performed on the resulting attenuation vs. frequency curve, with particular emphasis placed on testing for non-linearity.

11:30

3aUW11. Dependence of modal attenuation coefficient frequency variation on upper sediment attenuation. Wendy Saintval, William L. Siegmann (Dept. of Mathematical Sci., Rensselaer Polytech. Inst., 110 15th St., Troy, NY 12180), William M. Carey, Allan D. Pierce (Boston Univ., Boston, MA 02115), and James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

The range-averaged transmission loss increase in shallow water propagation depends critically on the intrinsic attenuation of the upper sediment. The attenuation coefficients of low-frequency (<1 kHz) propagating modes determine the frequency dependence. Ingenito [*J. Acoust. Soc. Am.* **53**, 858–863 (1973)] showed with measurements and theory that while individual mode attenuation coefficients decrease with frequency f , the sediment attenuation coefficient increases proportional to $f^{1.75}$. When results from many other shallow-water transmission experiments (broadband and narrowband) over sandy-silty sediments are compared to numerical calculations, it is found that a nonlinear-frequency dependent attenuation is required with an exponent between 1.5 and 2. The question considered here is how the intrinsic upper-sediment attenuation produces such behavior. A recent simplification of the Biot model [A. D. Pierce *et al.*, *J. Acoust. Soc. Am.* **114**, 2345 (2003)] has a power-law exponent of two. With this frequency-dependent bottom attenuation, a two-layer Pekeris waveguide yields modal attenuation coefficients that decrease with frequency as observed by Ingenito. However, a depth-dependent attenuation profile or a third near-surface layer with requisite properties can reverse this behavior. This suggests why higher-frequency numerical computations may require nonlinear frequency-dependent sediment profiles to calculate sound transmission accurately. [Work partially supported by ONR.]

11:45

3aUW12. Frequency dependence of the frame bulk moduli of granular marine sediments. Masao Kimura (Dept. of Geo-Environ. Technol., Tokai Univ., 3-20-1 Shimizu-Orido, Shizuoka, Shizuoka 424-8610, Japan, mkimura@scc.u-tokai.ac.jp)

The frame bulk modulus is important in analyzing the acoustic wave propagation in porous granular marine sediments. As the value of the frame bulk modulus, notable researchers in the community used about $10^8 P_a$ (constant value). However, the measured frequency characteristics

of the longitudinal wave velocity are not consistent with the calculated characteristics using the constant frame bulk modulus using the Biot-Stoll model. It was reported that the frame bulk moduli in water-saturated glass beads and beach sands are about $10^9 P_a$ at the frequency of 500 kHz, and are about ten times greater than that of air-saturated samples. It is also seen that the frame bulk moduli have dependence on the grain size. Moreover, these phenomena can be explained using the gap stiffness model [M.

Kimura, J. Acoust. Soc. Am. **116**, 2529 (2004)]. The grain size dependence of the frame bulk modulus suggests that there is the frequency dependence of the modulus. In this study, the measured results of the frequency characteristics of the longitudinal wave velocity in water-saturated glass beads and beach sand are compared with the calculated results using the Biot-Stoll model with the frequency dependent frame bulk modulus, derived using the gap stiffness model.

WEDNESDAY MORNING, 18 MAY 2005

STANLEY ROOM, 8:00 TO 9:30 A.M.

Meeting of Accredited Standards Committee (ASC) S2 Mechanical Vibration and Shock

to be held jointly with the

ANSI-Accredited U.S. Technical Advisory Group (TAG) Meetings for:
ISO/TC 108 Mechanical Vibration and Shock
ISO/TC 108/SC 2 Measurement and evaluation of mechanical vibration and shock as applied to machines, vehicles and structures
ISO/TC 108/SC 3 Use and calibration of vibration and shock measuring instruments
ISO/TC 108/SC 4 Human exposure to mechanical vibration and shock
ISO/TC 108/SC 5 Condition monitoring and diagnostics of machines and
ISO/TC 108/SC 6 Vibration and shock generating systems

R. J. Peppin, Chair S2
5012 Macon Road, Rockville, MD 20852

D. J. Evans, Vice Chair S2 and Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108 Mechanical Vibration and Shock
and Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC 3
Use and calibration of vibration and shock measuring devices
National Institute of Standards and Technology (NIST), 100 Bureau Drive, Stop 8221, Gaithersburg, MD 20899-8221

A. F. Kilcullen, Co-Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC 2
Measurement and evaluation of mechanical vibration and shock as applied to machines, vehicles and structures
5012 Woods Road, Hedgesville, WV 25427

R. Taddeo, Co-Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC 2
Measurement and evaluation of mechanical vibration and shock
as applied to machines, vehicles and structures
Naval Sea Systems Command 05T, 1339 Patterson Ave. SE, Stop 8800, Washington Navy Yard, Washington DC 20376-8800

D. D. Reynolds, Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC 4
Human exposure to mechanical vibration and shock
3939 Briar Crest Court, Las Vegas, NV 89120

D. J. Vendittis, Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC 5
Condition monitoring and diagnostics of machines
701 Northeast Harbour Terrace, Boca Raton, FL 33431

G. Booth, Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC 6
Vibration and shock generating systems
44 Bristol Street, Branford, CT 06405-4842

Accredited Standards Committee S2 on Mechanical Vibration and Shock. Working group chairs will report on the status of various shock and vibration standards currently under development. Consideration will be given to new standards that might be needed over the next few years. There will be a report on the interface of S2 activities with those of ISO/TC 108 and its subcommittees including plans for future meetings of ISO/TC 108 and/or its Subcommittees. The Technical Advisory Groups for ISO/TC 108 and its Subcommittees consists of members of S2 and other persons not necessarily members of those Committees. Open discussion of committee reports is encouraged.

Scope of S2: Standards, specifications, methods of measurement and test, and terminology in the field of mechanical vibration and shock, and condition monitoring and diagnostics of machines, including the affects of exposure to mechanical vibration and shock on humans, including those aspects which pertain to biological safety, tolerance and comfort.

WEDNESDAY MORNING, 18 MAY 2005

STANLEY ROOM, 10:30 A.M. TO 12:00 NOON

Meeting of Accredited Standards Committee (ASC) S3 Bioacoustics

R. F. Burkard, Chair S3

Hearing Research Laboratory, State University of New York at Buffalo, 215 Parker Hall, Buffalo, NY 14214

C. A. Champlin, Vice Chair S3

Department of Communication Sciences & Disorders, University of Texas, CMA 2-200, Austin, TX 78712

Accredited Standards Committee S3 on Bioacoustics. Working group chairs will report on the status of standards under development. Consideration will be given to new standards that might be needed over the next few years. Open discussion of committee reports is encouraged.

People interested in attending the meeting of the TAGs for ISO/TC 43 Acoustics and IEC/TC 29 Electroacoustics, take note - those meetings will be held in conjunction with the Standards Plenary meeting at 10:30 a.m. on Tuesday, 17 May 2005.

Scope of S3: Standards, specifications, methods of measurement and test, and terminology in the fields of psychological and physiological acoustics, including aspects of general acoustics, which pertain to biological safety, tolerance, and comfort.

WEDNESDAY AFTERNOON, 18 MAY 2005

REGENCY E, 1:00 TO 3:00 P.M.

Session 3pAA

Architectural Acoustics: Scattering in Architectural Acoustics

Kerrie G. Standlee, Chair

4900 SW Griffith Dr., Suite 216, Beaverton, OR 97005

Contributed Papers

1:00

3pAA1. Here and there, near and far: How proximity and separation affect scattering in concert halls. James B. Lee (6016 S. E. Mitchell, Portland, OR 97206)

Classical theory of scattering deals with plane waves, waves which are so far from their source that they form straight fronts of sound, with pressure amplitude and particle velocity exactly in phase. If such waves are much larger than objects they encounter inverse fourth power (Rayleigh) scattering obtains; if the waves are about the same size as the objects resonant (Mie) scattering occurs; if waves are much shorter than the objects specular (Ufimtsev) scattering is the rule. These all affect sound in the far field. But if sources are closer than a wavelength to objects their waves encounter the plane approximation is not valid; pressure amplitude and particle velocity are not in phase, so resonance-like phenomena occur. These occur on stages of concert halls: bass instruments producing waves 2 m or longer always are close to the floor; some, like tympani and violas, can be close to vertical surfaces too. This sort of scattering enhances fundamentals of notes with respect to the overtones, strongly affecting the timbre of such instruments.

1:15

3pAA2. Diagnosing scattering with the Wigner distribution. James B. Lee (6016 S. E. Mitchell, Portland, OR 97206)

The Wigner distribution resolves the power of a signal in both the time domain and the frequency domain. It also is a complete representation because it displays the relative phases of a signal's components as cross-products between them. This means that if one records the time-series of an impulse response in a room and computes a Wigner distribution from it, the kind of scattering, predominantly resonant or predominantly specular, is readily apparent, because the phase response of each is radically different. Specular scattering preserves the relative phases among components comprising the initial impulse and so preserves the impulse itself; resonant scattering randomizes the phases among components and so converts an impulse into a wide band of frequencies spread smoothly in time. Thus specular scattering tends to promote echoes and resonant scattering tends to promote diffusion. Graphic examples of each are given in the literature.

3pAA3. The influence of surface diffusivity on reverberation time prediction. Yan Zhang and Godfried Augenbroe (College of Architecture, Georgia Inst. of Technol., Atlanta, GA 30332)

It has been reported in the literature that the reverberation time tends to be overestimated if the ray-tracing simulation does not consider surface scattering effects. However no simple, quantitative method is available to estimate the effect of surface diffusivity on reverberation time. This article provides such a model to establish the relation between reverberation time and a surface diffusivity factor. It has also been observed that the influence of scattering reflections increases when the material is less uniformly distributed. Therefore the authors studied the current prediction method with non-uniformly distributed material and observed the following: Kuttruff equation based on the diffusive surface assumptions yields shorter reverberation times than Eyrings formula, while another equation containing reflective assumptions yields longer RT than Eyrings formula. It is demonstrated that the difference can reach up to 50%. Therefore the authors propose a linear combination of these two equations, with the proportion to each end determined by the surface diffusivity, assessed against Hann and Fricks method. A comparison between Boston Symphony Hall and New York Avery Fisher Hall is performed to demonstrate the significance of the surface diffusivity, and a Monte-Carlo simulation is included to support the result even under parameter uncertainties.

1:45

3pAA4. Aurally-adequate time-frequency analysis for scattered sound in auditoria. Molly K. Norris, Ning Xiang, and Mendel Kleiner (Dept. of Architectural Acoust., Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180)

The goal of this work was to apply an aurally-adequate time-frequency analysis technique to the analysis of sound scattering effects in auditoria. Time-frequency representations were developed as a motivated effort that takes into account binaural hearing, with a specific implementation of interaural cross-correlation process. A model of the human auditory system was implemented in the MATLAB platform based on two previous models [A. Härmä and K. Palomäki, HUTear, Espoo, Finland; and M. A. Akeroyd, A. Binaural Cross-correlogram Toolbox for MATLAB (2001), University of Sussex, Brighton]. These stages include proper frequency selectivity, the conversion of the mechanical motion of the basilar membrane to neural impulses, and binaural hearing effects. The model was then used in the analysis of room impulse responses with varying scattering characteristics. This paper discusses the analysis results using simulated and measured room impulse responses. [Work supported by the Frank H. and Eva B. Buck Foundation.]

2:00

3pAA5. Characteristics of scattered sound from overhead reflector arrays: Measurements at 1:8 and full scale. Timothy Gulsrud (Kirkegaard Assoc., 954 Pearl St., Boulder, CO 80302, tgulsrud@kirkegaard.com)

Measurements of sound scattering from reflector arrays have been carried out both at 1:8 scale in a controlled environment and at full scale in a concert hall. The results indicate wave interference between reflections from adjacent panels, resulting in an audible comb filter-like effect. Implications for the appropriate degree of overlap in coverage of adjacent panels will be discussed.

3pAA6. Effect of incidence angle and panel size on the phase of scattered energy from panel arrays. Jonathan Rathsam, Lily M. Wang (Architectural Eng. Program, Univ. of Nebraska Lincoln, 247 PKI, 1110 S. 67th St., Omaha, NE 68182-0681, jrathsam@mail.unomaha.edu), and Rendell R. Torres (Rensselaer Polytechnic Inst., Troy, NY 12180-3590)

This investigation focuses on factors that influence scattering from different ceiling panel arrays (e.g., incidence angle, and array size and density) by examining the scattered phase. A previous study by the third author [Torres and Vorländer, *Acta Acust.* (in press)] measured impulse responses using the maximum length sequence method at various receiver positions from scale model panel arrays of different sizes and densities. One interesting conclusion was that for multi-element arrays with panel dimensions and separation on the order of a wavelength, the reflected amplitude decreased, apparently due to destructive interference from inter-panel scattering. Analysis focused on the magnitude of the frequency response but the author noted that the phase curves could also contain information about scattered energy and the hypothesized inter-panel scattering. What the phase curves reveal is the subject of this study. It is expected that noticeable deviations from smoothness observed in the phase curves indicate non-specular scattering from the panels. The frequencies at which these deviations occur are examined to give insight into how the geometrical conditions of the panel array affect scattering.

2:30

3pAA7. A comparison of partially specular radiosity and ray tracing for room acoustics modeling. C. Walter BeamerIV (Dept. of Civil, Env., and Arch. Eng., Univ. of Colorado, Boulder, CO 80309) and Ralph T. Muehleisen (Illinois Inst. of Technol., Chicago, IL 60616)

Partially specular (PS) radiosity is an extended form of the general radiosity method. Acoustic radiosity is a form of bulk transfer of radiant acoustic energy. This bulk transfer is accomplished through a system of energy balance equations that relate the bulk energy transfer of each surface in the system to all other surfaces in the system. Until now acoustic radiosity has been limited to modeling only diffuse surface reflection. The new PS acoustic radiosity method can model all real surface types, diffuse, specular and everything in between. PS acoustic radiosity also models all real source types and distributions, not just point sources. The results of the PS acoustic radiosity method are compared to those of well known ray tracing programs. [Work supported by NSF.]

2:45

3pAA8. A new scattering method that combines roughness and diffraction effects. Claus Lyng Christensen (Odeon A/S c/o Acoust. Technol., rsted-DTU, Bldg. 352, DK-2800 Lyngby, Denmark)

Most of today's room acoustics programs make use of scattering coefficients which are used in order to describe surface scattering (roughness of material) and scattering of reflected sound caused by limited surface size (diffraction). A method which combines scattering caused by diffraction due to typical surface dimension, angle of incidence, and incident path length with surface scattering is presented. Each of the two scattering effects is modeled as frequency dependent functions. The benefits are two-fold: (a) Separating the user specified surface scattering coefficient from the room geometry, makes it easier for the user to make good guesses of the coefficients which will be in better agreement with the ones which can be measured. In many cases a scattering coefficient of say 1% for all surfaces may be sufficient. (b) Scattering due to diffraction is distance and angle dependent and as such it is not known before the actual raytracing or image source detection takes place. An example of this is that a desktop will provide a strong specular component to its user whereas it will provide scattered sound at far distances.

Session 3pAO**Acoustical Oceanography: Acoustical Oceanography Prize Lecture**

N. Ross Chapman, Chair

*Univ. of Victoria, School of Earth and Ocean Sciences, Victoria, BC V8W 3P6, Canada***Chair's Introduction—1:30*****Invited Paper*****1:35****3pAO1. Acoustic explorations of the upper ocean boundary layer.** Svein Vagle (OSAP, Inst. of Ocean Sci., P.O. Box 6000, Sidney, BC, Canada V8L 4B2, vagues@pac.dfo-mpo.gc.ca)

The upper ocean boundary layer is an important but difficult to probe part of the ocean. A better understanding of small scale processes at the air-sea interface, including the vertical transfer of gases, heat, mass and momentum, are crucial to improving our understanding of the coupling between atmosphere and ocean. Also, this part of the ocean contains a significant part of the total biomass at all trophic levels and is therefore of great interest to researchers in a range of different fields. Innovative measurement plays a critical role in developing our understanding of the processes involved in the boundary layer, and the availability of low-cost, compact, digital signal processors and sonar technology in self-contained and cabled configurations has led to a number of exciting developments. This talk summarizes some recent explorations of this dynamic boundary layer using both active and passive acoustics. The resonant behavior of upper ocean bubbles combined with single and multi-frequency broad band active and passive devices are now giving us invaluable information on air-sea gas transfer, estimation of biological production, marine mammal behavior, wind speed and precipitation, surface and internal waves, turbulence, and acoustic communication in the surf zone.

Session 3pID**Interdisciplinary: Hot Topics in Acoustics**

Beverly A. Wright, Chair

*Dept. of Communication Science and Disorders, Northwestern Univ., 2240 Campus Dr., Evanston, IL 60208***Chair's Introduction—1:30*****Invited Papers*****1:35****3pID1. The Student Council.** David T. Bradley (Univ. of Nebraska—Lincoln, 1110 S. 67th St., PKI 245 Omaha, NE 68182-0681, dbradley@mail.unomaha.edu)

Throughout its history, the Society has been a strong advocate for students exploring the world of acoustics. However, only recently was that support institutionalized and the responsibility placed in the hands of the students themselves. In the spring of 2000, a nascent body was formed to address the needs and concerns of Student Members of the ASA. The Vancouver meeting marks the 5th anniversary of this group, now formally known as the Student Council. The Council has evolved markedly in the past several years. The Council is currently composed of one student representative from each of the 13 technical committees and one Student Chapter Liaison. Recent student focused activities include the creation of a state-of-the-art website, the production of the first in a series of fellowship/grant workshops, and the presentation of a new award honoring the efforts of mentors in the Society. The Council's main focus is disseminating information of concern to students, advocating participation of students in Society activities and meetings, and promoting Society membership benefits to potential student members. Current and future programming and opportunities available to students, and those who work with students, will be presented and discussed.

1:55

3pID2. The ASA Regional Chapters program. Elizabeth McLaughlin (Chair of Committee on Regional Chapters) and Juan Arvelo (Co-Chair of Committee on Regional Chapters, Johns Hopkins Univ., Appl. Phys. Lab., 11100 Johns Hopkins Rd., Laurel, MD 20723-6099)

Are you involved in a Regional Chapter? The Regional Chapters Program certainly embraces the intent of our Society. "The ASA was founded . . . to increase and diffuse the knowledge of acoustics and promote its practical applications. Any person . . . interested in acoustics is eligible for membership." The history and the activities of each Chapter are unique. There are currently twenty active chapters uniquely positioned to promote acoustics through outreach and involvement with the public. There have been several new developments in the Regional Chapters Program, the most exciting being the incorporation of Student Chapters! Our first, the Nebraska Student Chapter, was approved at the 75th Meeting of the ASA in NYC. Several more are on their way! Existing Chapters are revitalizing! The Washington DC Chapter has recently found new enthusiasm, re-establishing a robust program. A new student scholarship has been organized thanks to the generosity of Larry and Julia Royster. Another recent enhancement is the expansion of the Regional Chapters Website. There one can find useful materials including an updated Chapter Start-up Kit. Involvement in a chapter is a great way to give back to the ASA, to learn, promote acoustics, to socialize, and to involve new persons in our exciting field!

2:15

3pID3. Imaging for understanding speech communication: Advances and challenges. Shrikanth Narayanan (Signal and Image Processing Inst., Univ. of Southern California, Viterbi School of Eng., Los Angeles, CA 90089-2564)

Research in speech communication has relied on a variety of instrumentation methods to illuminate details of speech production and perception. One longstanding challenge has been the ability to examine real-time changes in the shaping of the vocal tract; a goal that has been furthered by imaging techniques such as ultrasound, movement tracking, and magnetic resonance imaging. The spatial and temporal resolution afforded by these techniques, however, has limited the scope of the investigations that could be carried out. In this talk, we focus on some recent advances in magnetic resonance imaging that allow us to perform near real-time investigations on the dynamics of vocal tract shaping during speech. Examples include Demolin *et al.* (2000) (4–5 images/second, ultra-fast turbo spin echo) and Mady *et al.* (2001,2002) (8 images/second, T1 fast gradient echo). A recent study by Narayanan *et al.* (2004) that used a spiral readout scheme to accelerate image acquisition has allowed for image reconstruction rates of 24 images/second. While these developments offer exciting prospects, a number of challenges lie ahead, including: (1) improving image acquisition protocols, hardware for enhancing signal-to-noise ratio, and optimizing spatial sampling; (2) acquiring quality synchronized audio; and (3) analyzing and modeling image data including cross-modality registration. [Work supported by NIH and NSF.]

2:35

3pID4. Some recent condensed-matter applications of physical acoustics. Albert Migliori (Los Alamos Natl. Lab., MS E536, Los Alamos, NM 87545)

A non-inclusive sampling of recent advances in the application of physical acoustics to condensed matter physics is presented. Work by Ogi *et al.* on resonant ultrasound microscopy to map nano-scale variations in elastic moduli provides a new twist to the use of resonances. Laser-ultrasound studies by Wright and Matsuda on photoacoustic effects in ultra-thin metallic and semiconductor films leads to new meso-scale ultrasound studies, while Isaak and Ohno push the symmetry limits and temperature range of resonant ultrasound modulus studies. Migliori and Ledbetter extract the final pieces of an unusually-well-characterized charge-density-wave phase transition and examine very odd elastic behavior in materials with negative thermal expansion, while Pantea demonstrates an all-digital pulse echo system for high pressure work. Souslov probes hidden order with pulse-echo studies in pulsed magnetic fields. [Work supported by the National Science Foundation, the Department of Energy and the State of Florida.]

2:55

3pID5. Speech perception in the presence of other sounds. C. J. Darwin (Dept. of Psychol., Univ. of Sussex, Brighton BN1 9QG, UK)

The human listener's remarkable ability to recognize speech when it is mixed with other sounds presents a challenge both to models of speech perception and to approaches to speech recognition. This talk will review some of the work on how human listeners can perceive speech in sound mixtures and will try to indicate areas that might be particularly fruitful for future research.

3p WED. PM

Session 3pSA

Structural Acoustics and Vibration: Approximation Techniques for Fluid Loading Effects II

Joel M. Garrelick, Cochair

Cambridge Acoustical Associates/ETC, 84 Sherman St., Cambridge, MA 02140

David Feit, Cochair

*Applied Physical Sciences, Inc., 2 State St., New London, CT 06320***Invited Papers**

1:15

3pSA1. Doubly asymptotic approximations for transient and time-harmonic acoustics. Thomas L. Geers (Dept. of Mech. Eng., Univ. of Colorado, Boulder, CO 80309-0427)

The development and application of doubly asymptotic approximations (DAAs) in computational structural acoustics are reviewed. Transient radiation/scattering is emphasized, but application to time-harmonic radiation/scattering is also addressed. DAAs formulated for elastodynamic, poroelastic, and electromagnetic scattering are described that illustrate the breadth of application. The advantages and limitations of DAAs are described, and their relation to retarded-potential, similarity-based, and infinite-element methods are discussed.

1:45

3pSA2. Extension of infinite elements to explicit dynamic and eigenanalysis of submerged structures. Jeffrey Cipolla (ABAQUS, Inc., 166 Valley St., Providence, RI 02909-2499)

Infinite elements are distinct from alternative methods for exterior problems in that, like a finite element, they describe a small sub-region of the problem domain, and use locally supported shape functions to derive a method of weighted residuals statement thereon. Until recently, transient analyses using explicit time-integration methods, and eigenanalysis, were not possible using infinite elements. A formulation of acoustic infinite element compatible with implicit dynamics, time-harmonic acoustics, far-field extrapolation, explicit dynamics, and eigenanalysis is described. First steps in the development are the adoption of a basis corresponding to spherical radiating harmonics, the Bettess geometric map, and the Astley-Leis weighted residual formulation. A modified means to compute the element integrals, and a modification of the spherical harmonic basis, improve numerical conditioning of the element and stability of the formulation. The trivial frequency dependence of the Astley-Leis formulation, critical for its application to transient problems, also enables a formulation for eigenanalysis, which will be discussed. Finally, a splitting of the element operator into tangential mass- and stiffness-like terms enables a partial lumping of the element, and compatibility with explicit time integration schemes. Importantly, this splitting preserves the essential radiation mode accuracy of the formulation. Some computational examples are shown.

2:15

3pSA3. Numerical study of fluid loading approximations for submerged spherical shells using finite, infinite and boundary element methods. Rui M. Botelho (Electric Boat Corp., Dept. 462-M.S. J129-2, 75 Eastern Point Rd., Groton, CT 06340-4989, rbotelho@ebmail.gdeb.com), Michael J. Butler (Electric Boat Corp., Groton, CT 06340-4989), and John F. Waters (Electric Boat Corp., Groton, CT 06340-4989)

Numerical studies were conducted to compare acoustic fluid loading approximations for submerged spherical shells. Rigid and empty elastic shells submerged in a heavy acoustic fluid and subjected to harmonic excitation were considered. Solutions using boundary elements, infinite elements, and analytical approximations were compared, with special attention to modeling the near-field fluid beneath the infinite elements. In this case, the near-field fluid is typically modeled using fluid finite elements with frequency dependent mesh refinement and radial extent. The spherical shell is modeled in the same manner using shell finite elements for both boundary and infinite element solutions in order to provide a direct comparison of the fluid loading approximations. Convergence studies involving the near-field fluid were performed with emphasis on the effects of meshing parameters, including refinement, mesh grading, and polynomial order of the radial basis functions for p -version fluid finite elements. Results of interest include comparisons of the submerged structural response, focusing on the reactive and resistance parts of the fluid impedance.

Session 3pUW

Underwater Acoustics and Acoustical Oceanography: Frequency Dependence of Sound Speed and Attenuation in Marine Sediments II

Gopu Potty, Cochair

Dept. of Ocean Engineering, Univ. of Rhode Island, Narragansett, RI 02882

Peter H. Dahl, Cochair

Applied Physics Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698

Contributed Papers

1:30

3pUW1. Frequency dependence of elliptical particle motion of acoustic waves transmitted into the seabed from a point source in water. David M. F. Chapman, Paul C. Hines, and John C. Osler (Defence Res. and Development Canada Atlantic, P.O. Box 1012, Dartmouth, NS, Canada B2Y 3Z7, dave.chapman@drdc-rddc.gc.ca)

When perpendicular components of the particle velocity of a continuous acoustic wave have unequal amplitude and phase, the displacement traces an elliptical path. This is caused by inhomogeneous waves and/or losses in the medium. For spherical waves radiating from a point source in water and transmitted into the seabed, the orientation of the elliptical orbit varies with receiver location and frequency, and the ellipse major axis is not always parallel to the direction of propagation suggested by Snell's law. If not accounted for, this could bias estimates of sediment sound speed derived from particle velocity measurements. This presentation describes a simple orbit model based on numerical evaluation of integrals for the transmitted field over the entire wavenumber range. The model provides rapid simulation of results for different experimental geometries and source frequencies. The model is validated using an established full-field model and compared with experimental data. [Work supported in part by ONR Code 32.]

1:45

3pUW2. Acoustic velocity measurements in seafloor sands at frequencies from 1 to 400 kHz. Michael A. Zimmer, L. Dale Bibee, and Michael D. Richardson (Naval Res. Lab., 1005 Balch Blvd., Stennis Space Center, MS 39529-5004)

We measured the acoustic velocity and attenuation at frequencies from 1 kHz to 400 kHz in shallow seafloor sands off Fort Walton Beach, FL, USA. We used three separate systems with overlapping frequency ranges in order to cover this entire frequency range. For frequencies from 1 to 20 kHz, we implanted a seafloor array of 35 hydrophones and 5 three-component accelerometers at depths from 0 to 1 m over a 4 m by 4 m area, and recorded signals from two acoustic sources positioned at offsets from 1 m to 20 m. Measurements from 15 to 120 kHz were made at 30-cm sediment depths with the *In Situ* Sediment Acoustic Measurement System (ISSAMS), which consists of a linear array of 4 piezoelectric probes; the outer probes transmit a single-frequency burst while the inner probes act as receivers. Additionally, velocity measurements were made on diver-collected cores (5–20 cm sediment depths) at frequencies from 50 to 400 kHz. We present comparisons of the measured frequency dependence of the acoustic velocity and attenuation to theoretical predictions from various constitutive and wave propagation models, parameterized with inputs from detailed characterizations of diver-collected cores.

2:00

3pUW3. Inverting for sandy sediment sound speed in very shallow water using boat noise. Dajun Tang (Appl. Phys. Lab, Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

Small boat propeller noise was recorded on a single hydrophone in very shallow water for the purpose of estimating sediment sound speed. The experiment was conducted near shore in water depth between 2 and 10 m. The sediment consisted of uniform sand. The lone hydrophone was moored 2 m above the bottom at 6 m depth. A small boat traveling at constant speed was used as the sound source, and ran both parallel and perpendicular to shore. Thus, both the range independent waveguide case and wedge shaped waveguide case could be investigated. The source tracks were recorded by using a GPS recorder on the boat. Water depth in the entire area was measured, as was the sound speed profile at the receiver. The processed data sets resulted in interference patterns in range-frequency plots. Aided by numerical simulations, sediment sound speed can be estimated over the frequency range of 500–4500 Hz. [Work supported by the U.S. Office of Naval Research.]

2:15

3pUW4. Effect of frequency dependence of complex sound speed on impulse response in littoral seas. David Knobles (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029)

For certain classes of sediments common to littoral ocean environments, there exists theoretical reasons and experimental evidence that the real part of the sound speed is frequency dependent. Further it has been observed for more than 30 years that to model transmission loss data below 1 kHz for such sediments, it is often required that the attenuation possess a non-linear frequency dependence. However, analyses of broadband propagation below 2 kHz in littoral waveguides commonly assume that sound speeds in marine sediments are frequency independent and that the attenuation depends linearly on frequency. It is natural to explore how the neglect of the frequency dependence of the complex sound speed affects the predictability of broadband propagation and under what conditions can inversion methods estimate the functional form of the frequency dependence. In this study simulated impulse responses for typical littoral waveguides over the 10–2000 Hz band are generated assuming a Biot-Stoll formulation. The effects of the frequency dependence of the complex sound speed on the impulse response are examined within the context of how does one form a valid inverse problem that would allow for the frequency dependence of the complex sound speed to be uniquely determined. [Work supported by ONR.]

2:30

3pUW5. Longitudinal wave dispersion and attenuation in a two phase granular material. Hasson Tavossi (Dept. of Physical and Environ. Sci., Mesa State College, 1100 North Ave., Grand Junction, CO 81501)

Longitudinal elastic wave-velocity and attenuation in a two phase granular material are investigated, in the cases of a random arrangement of grains in contact with each other, surrounded by water or air. The velocity dispersion and attenuation of a longitudinal wave are measured experimentally, at low and high frequencies and are analyzed in light of a modified

Biot theory. The results of these analyses on wave velocity and attenuation are expressed as a function of frequency, grain size, porosity, viscosity of the fluid phase, as well as elastic moduli and densities of the solid and fluid phases. It is shown that longitudinal wave velocity in such a material decreases with grain size and frequency at high frequencies. Results obtained by increasing the grain size resemble those measured by increasing the frequency. Wave attenuation is expressed as a function of frequency, porosity, grain size and fluid viscosity. Theoretical results are compared with the experimental findings on the wave velocity and attenuation.

WEDNESDAY AFTERNOON, 18 MAY 2005

REGENCY A, B, C, 3:30 TO 5:30 P.M.

Plenary Session, Business Meeting and Awards Ceremony

William A. Kuperman, President
Acoustical Society of America

Stan E. Dosso, President
Canadian Acoustical Association

Welcome Remarks

Annual Business Meeting of the Acoustical Society of America

Presentation of Certificates to New Fellows

Catherine T. Best
Bennett M. Brooks
Michael G. Brown
Charles F. Gaumont
Murray R. Hodgson

Presentation of Acoustical Society Awards

Medwin Prize in Acoustical Oceanography to Svein Vagle
Silver Medal in Psychological and Physiological Acoustics to H. Steven Colburn
R. Bruce Lindsay Award to Lily M. Wang
Helmholtz-Rayleigh Interdisciplinary Silver Medal to Gilles A. Daigle
Gold Medal to Allan D. Pierce

Session 3eAA**Architectural Acoustics, Education in Acoustics, Noise and Psychological and Physiological Acoustics:
Topical Meeting on Classroom Acoustics—Public Forum**

Murray R. Hodgson, Cochair

*Univ. of British Columbia, School of Occupational and Environmental Hygiene, 2206 East Mall,
Vancouver, BC V6T 1Z3, Canada*

Maureen MacDonald, Cochair

*Elementary School Teacher, Vancouver, BC, Canada***Chair's Introduction—6:30****6:35****3eAA1. The need for good acoustic design of schools.** Bridget Shield (Dept. of Eng. Systems, Faculty of Eng. Sci. and Built Environment, London South Bank Univ., London SE1 0AA, UK)

This paper gives an overview of research into classroom acoustics, highlighting the importance of a good acoustic environment in schools to enhance teaching and learning. The paper is aimed at a general audience of people interested in education and school design. In the past 30 years there has been a great deal of research into the effects of noise and poor acoustics in schools on children and teachers. It has been shown in many studies that children have difficulty hearing and understanding their teachers in noise, and both external environmental noise and noise within a school affect children's academic performance. Furthermore many teachers suffer from voice and throat problems which may be attributable to a poor acoustic environment in the classroom. The acoustic design of a classroom has a direct influence upon noise levels and the intelligibility of speech. Poor sound insulation and excessive reverberation have the potential to increase noise levels and reduce speech intelligibility. However, despite the introduction in many countries of legislation or guidelines for acoustic design of schools, in general acoustics still has a low priority in school design and many schools, old and new, fail to meet the current standards.

7:35–9:35**Panel Discussion**

ACOUSTICAL SOCIETY OF AMERICA
Silver Medal in
Psychological and Physiological Acoustics



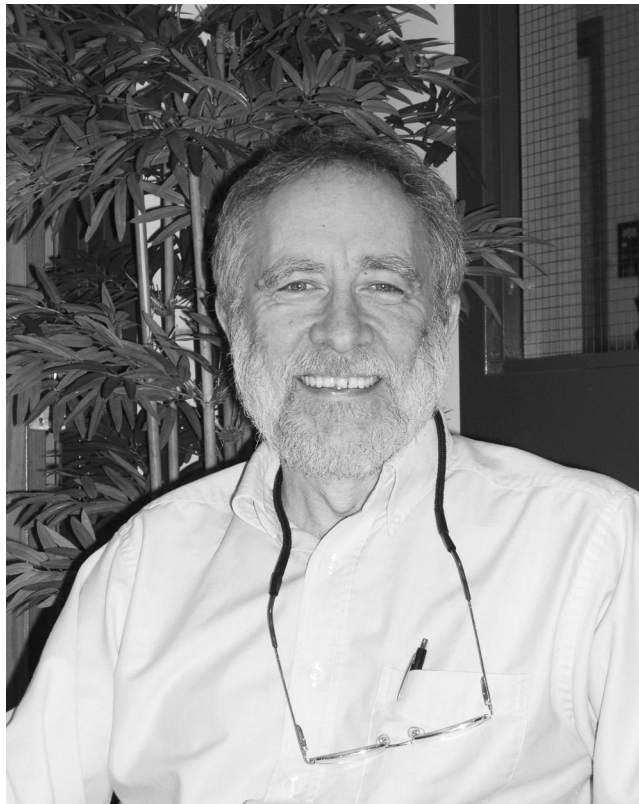
H. Steven Colburn

2004

The Silver Medal is presented to individuals, without age limitation, for contributions to the advancement of science, engineering, or human welfare through the application of acoustic principles, or through research accomplishment in acoustics.

PREVIOUS RECIPIENTS

Lloyd A. Jeffress	1977
Ernest Glen Wever	1981
Eberhard Zwicker	1987
David M. Green	1990
Nathaniel I. Durlach	1994
Neal F. Viemeister	2001
Brian C. J. Moore	2002



ENCOMIUM FOR H. STEVEN COLBURN

... for contributions to psychological and physiological aspects of binaural hearing

VANCOUVER, CANADA • 18 MAY 2005

Steve Colburn was born in Portsmouth, Ohio on June 23rd, 1941. He attended Clay Township Rural High School in a farming community where his main interest was playing trumpet in the high school band and in a "dance band". He moved to Boston in 1959 to attend the Massachusetts Institute of Technology (MIT), being only the second person from his high school to attend college outside Ohio. He has lived and worked in Boston ever since.

At MIT Steve received S.B., S.M., and Ph.D. degrees during the 1960's and served as a faculty member in the Electrical Engineering Department for a number of years thereafter. From the early 1980's to the present, he has been a faculty member in the Biomedical Engineering Department at Boston University (BU), serving as Chair of the Department for about a decade. Steve's extended family includes, in addition to his wife Theo, three sons, a daughter-in-law, and two grandchildren.

Steve's scientific contributions in the area of binaural hearing, his main area of interest, are unique in their overall breadth; the extent to which they integrate perceptual data, physiological data, and sophisticated mathematical modeling; and, more generally, their profound influence on the criteria employed to define what it means to be a serious auditory scientist.

The breadth of his research is evident in his work published in *The Journal of the Acoustical Society of America* (JASA) over the past 40 years on the relation of binaural detection to lateralization, the effects of envelope and carrier interaural time delays, the implications of frozen-noise detection experiments for models of binaural detection, the roles played by peripheral versus central limitations on perceptual performance, the challenges faced by the binaural system in complex acoustic environments, and a large variety of issues associated with hearing impairments and hearing aids (including cochlear implants). Also noteworthy are the two review chapters in *The Handbook of Perception* on binaural hearing that he coauthored in 1979 and that still serve as major sources of information on both empirical and theoretical results underlying our understanding of binaural hearing. Finally, and perhaps most importantly, Steve's elaboration of Jeffress' model of binaural interaction, which appears in various forms in many of his JASA articles, has not only contributed greatly to binaural hearing theory, but also constitutes a wonderful tribute to perhaps the most highly revered figure in this field (and the first Silver Medal recipient in Psychological and Physiological Acoustics).

Steve's remarkable ability to exploit both experimental data and mathematical-modeling tools to gain deep understanding of the auditory system is clearly evidenced by his research on the performance that would be achieved by an ideal processor operating on the noisy outputs of the peripheral processing observed in the auditory system. In general, this approach provides a "synthetic surgical knife" to the problem of determining where in the auditory system the loss of information about an acoustic stimulus is most severe. To the extent that performance of this model matches the perceptual performance observed in humans, the major loss is in the periphery. To the extent that the performance of the model exceeds that of the human, the major loss is central. Steve's exploitation of this approach in the area of binaural hearing began while he was working with William Siebert at MIT in the 60's and has continued to this day. As time passes and quantitative physiological knowledge of neural activity includes increasingly central activity, the points at which this "synthetic surgical knife" can be applied include increasingly central locations in the auditory system.

In addition to Steve's outstanding direct contributions to auditory science, he has made enormous indirect contributions via his teaching and administrative/organizational work. His devotion to students and his excellence as a teacher are evidenced by his superior work as housemaster while at MIT, the many teaching awards he has received at BU, the great warmth shown to him by all of his students, and the substantial contributions these students have themselves made to auditory science. His great skills as an administrator/organizer have been clearly revealed by his work as Department Chair at BU; his creation and leadership of the Hearing Research Center at BU (which is recognized internationally as a leading center in auditory science); his contributions as an Associate Editor of *The Journal of the Acoustical Society of America*; and his service as a member of various review committees of the National Institutes of Health. Most remarkable,

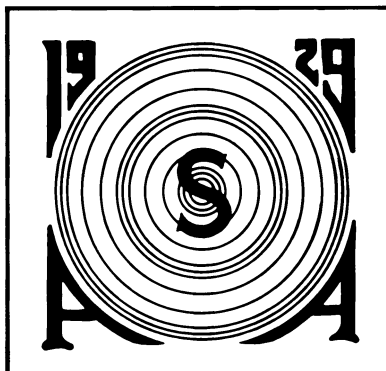
perhaps, is his creation and leadership of the “Binaural Bash” at BU, an annual meeting dedicated to informal exchange of ideas, problems, and results, which is attended by a wide variety of scientists from all over the world.

As evidenced early on by his exploits as a mountain climber and cold-weather camper (who used to enjoy sleeping in a tent on Mt. Washington at -30° with 40 mph winds), Steve has a long history of addressing and solving challenging problems of many kinds.

NATHANIEL I. DURLACH
CONSTANTINE TRAHOTIS

ACOUSTICAL SOCIETY OF AMERICA

R. BRUCE LINDSAY AWARD



Lily M. Wang

2005

The R. Bruce Lindsay Award (formerly the Biennial Award) is presented in the Spring to a member of the Society who is under 35 years of age on 1 January of the year of the Award and who, during a period of two or more years immediately preceding the award, has been active in the affairs of the Society and has contributed substantially, through published papers, to the advancement of theoretical or applied acoustics, or both. The award was presented biennially until 1986. It is now an annual award.

PREVIOUS RECIPIENTS

Richard H. Bolt	1942	Peter N. Mikhalevsky	1984
Leo L. Beranek	1944	William E. Cooper	1986
Vincent Salmon	1946	Ilene J. Busch-Vishniac	1987
Isadore Rudnick	1948	Gilles A. Daigle	1988
J. C. R. Licklider	1950	Mark F. Hamilton	1989
Osman K. Mawardi	1952	Thomas J. Hofler	1990
Uno Ingard	1954	Yves H. Berthelot	1991
Ernest Yeager	1956	Joseph M. Cuschieri	1991
Ira J. Hirsh	1956	Anthony A. Atchley	1992
Bruce P. Bogert	1958	Michael D. Collins	1993
Ira Dyer	1960	Robert P. Carlyon	1994
Alan Powell	1962	Beverly A. Wright	1995
Tony F. W. Embleton	1964	Victor W. Sparrow	1996
David M. Green	1966	D. Keith Wilson	1997
Emmanuel P. Papadakis	1968	Robert L. Clark	1998
Logan E. Hargrove	1970	Paul E. Barbone	1999
Robert D. Finch	1972	Robin O. Cleveland	2000
Lawrence R. Rabiner	1974	Andrew J. Oxenham	2001
Robert E. Apfel	1976	James J. Finneran	2002
Henry E. Bass	1978	Thomas J. Royston	2002
Peter H. Rogers	1980	Dani Byrd	2003
Ralph N. Baer	1982	Michael R. Bailey	2004



CITATION FOR LILY M. WANG

... for contributions to room and musical acoustics

VANCOUVER, CANADA • 18 MAY 2005

Lily Wang has a contagious enthusiasm for life and acoustics, reinforced with responsibility, determination and resilience. Both of her parents were born in China and were raised in Taiwan. After her father received his degrees, her family moved to Chattanooga, Tennessee where he worked for the Tennessee Valley Authority. At the age of 11, Lily was involved in an automobile accident that took her mother's life. She continues to motivate herself with her mother's image in the back of her mind. While in high school she read about careers in architectural acoustics. This led her to a Bachelor of Science degree in Civil Engineering (with a Certificate in Architecture) from Princeton University, a Ph.D. in Acoustics from Penn State, an ASA Hunt Fellowship and a faculty position in architectural acoustics at the University of Nebraska.

She graduated from an all-female high school as class Valedictorian and was selected to be a United States Presidential Scholar, a program established to recognize and honor the nation's most distinguished graduating high school students. On her way to graduating from Princeton Magna Cum Laude, Lily began her foray into acoustics with summer internships at Georgia Tech and Jaffe Holden Associates, studying the integration of acoustics, aesthetics and technical design of fabric tension membrane structures.

Lily won National Science Foundation (NSF) and American Association of University Women (AAUW) fellowships to support her graduate studies in acoustics at Penn State. At Penn State, her interest in music, acquired at the age of four, led to her study of the acoustic radiation from violins. She recognized that there are three elements in the sound production from violins; excitation by the bow/string interaction, vibration of the violin body and acoustic radiation from the body, and that the last element had received far less attention than the first two elements. However, the analysis of the acoustic radiation from violins: required visualization of the acoustic field with a spatial resolution smaller than an acoustic wavelength. This required the application of nearfield acoustical holography to a violin played continuously over long periods of time. Lily applied ingenuity and tenacity in designing and building a bowing machine to excite violins and a nearfield acoustical holography system that surrounded the bowed violin, within the budgetary constraints typical of research on the acoustics of musical instruments. Her dissertation results were received with great interest by the musical acoustics community. She was invited three times to present papers at ASA meetings on her research. Her Ph.D. thesis led to several publications, one in *The Journal of the Acoustical Society of America*. Carleen Hutchins (and people passing by during a demonstration in Carleen's driveway) were impressed with the bowing machine. Carleen was also impressed with Lily, turning to me to tell me that she thought Lily knew more than I did. She got no argument there. The bowing machine now hangs from the rafters of my garage. Lily now says that she loved her time at Penn State, but only after the pain of sewing hundreds of horse hairs into the belt used in the bowing machine has diminished with time.

In the midst of her Ph.D. studies, Lily worked one summer at Bell Laboratories with Gary Elko on the application of boundary layer models to the turbulent field of a talker's mouth during speech plosives, and to methods of attenuation of the effects of the turbulent field. She demonstrated her ability to mix theory and measurements in developing a simple model that produced predictions that agreed with measured results.

Upon completion of her Ph.D., Lily was awarded the ASA F.V. Hunt Postdoctoral Research Fellowship in Acoustics. She then returned to her first interest, architectural acoustics, to study the subjective responses to acoustic fields in rooms, including auralization, at the Technical University of Denmark with Anders Christian Gade and Jens Holger Rindel. Her work was valuable and she has been invited back to the Technical University of Denmark to continue working on collecting data that relate subjective responses and objective measures of acoustic fields in large spaces, one of the most challenging problems in architectural acoustics. When Lily arrived at the University of Nebraska, the Architectural Engineering Department Program in the Peter Kiewit Institute had no acoustics program. In less than five years, she has developed one of the best programs in architectural acoustics in the country. She has transferred her enthusiasm to her students, many of whom are now active in the Society. Her work on spatial impression of musical acoustical fields in auditoria, objective measures in concert halls, effects of sound directivity and scattering on auralization, computer modeling of acoustic fields in rooms, classroom acoustics, reverberation and acoustic fields in coupled spaces, correlation of subjective and objective measures of speech

intelligibility, and the combined effects of temperature and noise on human performance are well known and favorably received as indicated by the seven invited ASA papers on these research projects. Lily is now one of the 'stars' in architectural acoustics and her impact on building a strong program at the University of Nebraska has been exemplary.

She is an active and popular contributor to the Society. In addition to being invited ten times to present papers at ASA meetings, she was one of the nine young presenters selected to speak on the future of acoustics at the Society's 75th anniversary meeting in May 2004. Lily is now the chair of the Architectural Acoustics technical committee, one of the largest and most active (and sometimes unwieldy) of the ASA technical committees. She is a member of the Newman Student Award Fund Advisory Committee and the Committee on Women in Acoustics, and has worked hard to develop the flourishing Architectural Acoustics Student Design competition.

In addition to her outstanding technical and educational contributions, Lily is a pleasure to be with. She is always surrounded by people at ASA meetings. I am sure we will see a lot of Lily in the future as she continues to serve the Society, the academic community and the world of acoustics with distinction.

Lily lives with Oliver, who I am sure, is one spoiled cat.

COURTNEY B. BURROUGHS

ACOUSTICAL SOCIETY OF AMERICA
Helmholtz-Rayleigh Interdisciplinary
Silver Medal
in
Noise and Physical Acoustics



Gilles A. Daigle
2005

The Silver Medal is presented to individuals, without age limitation, for contributions to the advancement of science, engineering, or human welfare through the application of acoustic principles, or through research accomplishment in acoustics.



CITATION FOR GILLES A. DAIGLE

... for contributions to understanding the effects of micrometeorology, topography, and ground properties on outdoor sound propagation

VANCOUVER, CANADA • 18 MAY 2005

Gilles Daigle received a B.Sc. in physics from the Université de Moncton, New Brunswick, in 1975. He intended to do graduate work in nuclear physics, but that summer he worked as a student in the Acoustics Section at the National Research Council (NRC) in Ottawa and Gilles' career path underwent an abrupt change. He decided he would prefer to pursue a life in acoustics instead. NRC is not a degree-granting institution and so special arrangements were made with Carleton University in Ottawa. Gilles did the required coursework at the university and the acoustical work, both theoretical and experimental at NRC under the supervision of Tony Embleton and Joe Piercy.

Between 1975 and 1981 Gilles' work was devoted mainly to measuring the meteorological parameters of wind and temperature gradients, and the spectrum of sizes and intensity of turbulence. He related these parameters theoretically and by measurement, under many different weather conditions, to the resulting fluctuations of sound pressures. In this period, the titles of his M.Sc. and Ph.D. theses were "Effects of atmospheric turbulence on the interference of acoustic waves due to the presence of a boundary" and "Interference, diffraction and scattering of acoustic waves in a turbulent atmosphere." Apart from new theory, this work developed new measuring techniques using very short-time-constant electrical thermoanemometers originally developed for the air-conditioning and heating industry.

Since 1981 Gilles and his colleagues both at the National Research Council and elsewhere have systematically measured, and compared with the new theories, each of the many mechanisms that are relevant to sound propagation outdoors. These mechanisms include interference between direct waves and those reflected at the ground surface (often at near-grazing incidence where ground waves and surface waves can play a dominant role), the acoustical properties of many types of ground, diffraction over barriers, penetration into shadow zones, scattering by atmospheric turbulence, and refraction by vertical gradients of wind and temperature. Also included in Gilles' range of studies are more complicated, but realistic, ground surfaces such as layers of snow over harder ground surfaces and propagation across an asphalt-to-grass interface. Gilles' contributions to theoretical understanding, to experimental work indoors, and to field measurements outdoors have been substantial and sometimes unique. He has provided a theoretical structure that unifies several previously-incompatible databases and explains their deviations from traditional predictions. These databases were derived either from drive-by tests of motor vehicles at ranges up to 50 feet, from community noise measurements (50 to 1000 feet from the sources), or from measured aircraft or airport noise levels at ranges greater than a thousand feet. Often acousticians elsewhere have studied parts of this technical area in greater depth but less breadth than Gilles'; but he has worked in all the areas and so has a broad overview of the subject.

Gilles has shown that meteorological and topographical effects are noticeable and usually dominant for all distances of propagation greater than a few feet. This work is laying the foundation for community noise prediction schemes that can accurately predict sound levels for specific rather than average situations. These include different daytime and nighttime atmospheric profiles, and whether the listener is on the ground, in a highrise building, up a hillside or behind a hill. The economic benefits of this more accurate knowledge are immense—consider the economic consequences of decisions about the siting of airports and major highways, and the social benefits of locating highway noise barriers where they will provide the expected acoustical performance.

The peer recognition accorded Gilles' technical work is worldwide. The recognition started early when he received the Governor General's Gold Medal in 1975 from the Université de Moncton (awarded annually to the top undergraduate student, one in each Canadian university). He received the R. B. Lindsay Award of the Acoustical Society in 1988 and was a Plenary Session speaker at the 15th International Congress of Acoustics in Norway in 1995. He was a Guest Researcher at the Agency of Industrial Science and Technology in Japan in 1996, was awarded the Médaille Etrangère by the Société Française d'Acoustique in 2002, and has co-supervised over a dozen M.Sc. and Ph.D. students in both Canadian and French universities.

Gilles serves the Acoustical Society with great energy and enthusiasm. He has been chair of the Technical Committee on Physical Acoustics (1987-1990), chair of the Committee on Meetings (1996-1999), Associate Editor of *The Journal of the Acoustical Society of America* for Noise and its Effects and Control (1993-1997), chair of the Ottawa meeting in 1993, member of the Medals

and Awards Committee (1991-1994) and of the Nominating Committee for the 1993 election. Gilles was elected as a member of the Executive Council (1992-1995), as Vice-President Elect and Vice President he served as an Officer of the Society from 1999 to 2001, and while Past Vice President for one more year as a member of the Executive Council. Service to other acoustical organizations includes serving as Secretary General of the International Commission on Acoustics (1998-2001) and as its President (2001-2004). For the International Institute of Noise Control Engineering, he chaired an international study group on the effectiveness of roadside noise barriers, was Vice President for Technical Initiatives (1996-2002) and is currently its President Elect. Gilles has served different levels of government in a variety of ways, from training courses for local police forces to being a technical expert on a federal government environmental panel on military flying activities in Labrador and Quebec.

Gilles Daigle is a leader in the field of outdoor sound propagation with knowledge that spreads beyond acoustics to encompass the related factors of topography and micrometeorology. He serves as a leader of the international acoustical community through several of its societies and organizations. It is fitting that today the Acoustical Society of America recognizes Gilles Daigle by its most appropriate award, the Helmholtz-Rayleigh Interdisciplinary Silver Medal in Noise and Physical Acoustics.

TONY F.W. EMBLETON
MICHAEL R. STINSON

GOLD MEDAL of the Acoustical Society of America



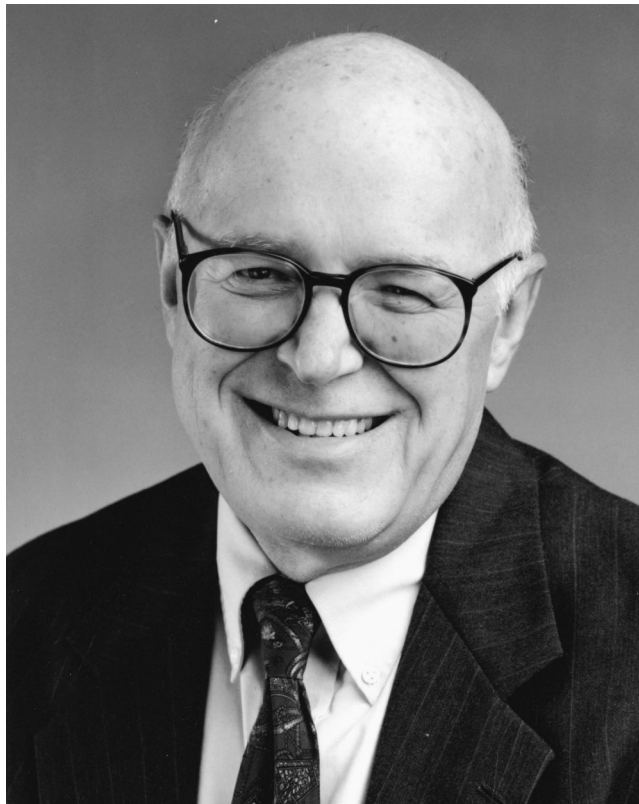
Allan D. Pierce

2005

The Gold Medal is presented in the spring to a member of the Society, without age limitation, for contributions to acoustics. The first Gold Medal was presented in 1954 on the occasion of the Society's Twenty-Fifth Anniversary Celebration and biennially until 1981. It is now an annual award.

PREVIOUS RECIPIENTS

Wallace Waterfall	1954	Cyril M. Harris	1987
Floyd A. Firestone	1955	Arthur H. Benade	1988
Harvey Fletcher	1957	Richard K. Cook	1988
Edward C. Wentz	1959	Lothar W. Cremer	1989
Georg von Békésy	1961	Eugen J. Skudrzyk	1990
R. Bruce Lindsay	1963	Manfred R. Schroeder	1991
Hallowell Davis	1965	Ira J. Hirsh	1992
Vern O. Knudsen	1967	David T. Blackstock	1993
Frederick V. Hunt	1969	David M. Green	1994
Warren P. Mason	1971	Kenneth N. Stevens	1995
Philip M. Morse	1973	Ira Dyer	1996
Leo L. Beranek	1975	K. Uno Ingard	1997
Raymond W. B. Stephens	1977	Floyd Dunn	1998
Richard H. Bolt	1979	Henning E. von Gierke	1999
Harry F. Olson	1981	Murray Strasberg	2000
Isadore Rudnick	1982	Herman Medwin	2001
Martin Greenspan	1983	Robert E. Apfel	2002
Robert T. Beyer	1984	Tony F. W. Embleton	2002
Laurence Batchelder	1985	Richard H. Lyon	2003
James L. Flanagan	1986	Chester M. McKinney	2004



CITATION FOR ALLAN D. PIERCE

. . . for contributions to physical, environmental, and structural acoustics, acoustics education, and leadership as Editor-in-Chief of the Society.

VANCOUVER, CANADA • 18 MAY 2005

Iowa has made important contributions to acoustics; one of them is Allan Pierce, who was born December 18, 1936 in Clarinda. Allan's family lived in Wichita, Kansas during World War Two, then moved to Las Cruces, New Mexico, where he attended New Mexico State University. Two and a half years were sufficient for him to earn a B.S. with highest honors in Physics in 1957. From there, he pursued his graduate studies in the Physics Department of the Massachusetts Institute of Technology, which marks the beginning of his well-known affinity for the Boston area. In 1962 he was awarded the Ph.D. in physics for his thesis on electron lattice interactions in the Born-Oppenheimer approximation. With the encouragement of Philip Morse, who was on his Ph.D. committee, Allan went to work for the Rand Corporation in Santa Monica. We are thankful that Allan was new to the company. Otherwise Albert Latter probably would not have turned to Allan when the senior researchers working in quantum physics turned down a project in atmospheric propagation, thereby diverting Allan from the area of his doctoral work to acoustics.

Allan's second stay in the Boston area came when he transferred to the Avco Corporation in Wilmington from 1963 to 1966. He then returned to MIT as an Assistant Professor, where he was promoted to Associate Professor in 1968. Allan departed for the warmer climate of Atlanta in 1973 when he moved to the Georgia Institute of Technology as a Professor of Mechanical Engineering. He was awarded a Regents' Professorship there in 1976. During his tenure at Georgia Tech he had visiting positions at the Max Planck Institut für Strömungsforschung from 1976 to 1977, and the Transportation Systems Center of the U.S. Department of Transportation, which was his third stay in Boston. He departed Georgia Tech in 1988 for Pennsylvania State University, where he was a Professor and the Leonhard Chair Holder, with joint appointments in the Department of Mechanical Engineering and the Graduate Program in Acoustics. His most recent return to the Boston area was in 1993, when he became the Chair of the Department of Aerospace and Mechanical Engineering at Boston University. He stepped down from that position in 1999, but continues there as a Professor. Throughout his career, his wife Penelope Claffey, who he married in 1961, has been by his side. She and their children, Jennifer and Bradford, are justifiably proud of his accomplishments. He has successfully managed the difficult task of being a devoted husband and father, while simultaneously being fully committed to his profession.

Allan is an intellectual in the truest sense. If you cannot find Allan during a professional meeting, look for him in an antique bookstore. His interests are universal, and his research in a diverse range of topics always blends physical insight and mathematical rigor. Allan's early work addressed propagation of sound in the atmosphere. In a Letter to the Editor in *Nature*, he described how the magnitude of an explosion could be determined from its measured waveform, which was used to verify the ban on atmospheric testing of nuclear weapons, and subsequently to estimate the magnitude of the Mt. Saint Helen's eruption. His analysis explaining the effects of turbulence on sonic boom rise time remains one of the best available theories. It was a logical transition for him when he turned to problems in propagation in natural waveguides. The theory he developed to predict propagation in natural waveguides whose properties vary slowly with horizontal distance is still quoted today. Interestingly, this work was guided by analogies he drew with his thesis work. In fact, Allan's ability to recognize analogies between apparently different areas is an important attribute of his approach.

A complete survey of Allan's work would need to describe his contribution to development of the surface variational principle for radiation and scattering, and his application of geometrical acoustics concepts to describe wave propagation of plates and shells, which led to new insights into the interpretation of experimental data. A hot topic several years ago was "fuzzy structures," about which he provided mathematical insight to a much debated subject. One should also recognize his efforts in laser generation of sound, and diffraction around thick barriers. In each area, Allan's papers are at the core, and often the first to be cited by others.

Research is only one aspect of Allan Pierce. He was the chief architect of two outstanding programs in acoustics. At Georgia Tech, the prospect of working with him enticed several individuals, including me, to join the School of Mechanical Engineering. As department chair at Boston University he convinced the administration to expand the faculty, with many of the subsequently recruited faculty having research interests in acoustics. At Penn State, where the acous-

tics program was highly respected prior to his arrival, Allan provided an intellectual focus as a faculty colleague, as well as the Technical Director of the Center for Acoustics and Vibrations. In every academic position Allan has held, he has had a profound impact on a generation of young acousticians who he nurtured and mentored, and now are highly accomplished members of our Society. An important aspect of Allan's educational activities is his textbook, *Acoustics: An Introduction to Its Physical Principles and Applications*, which is one of the best-sellers in the ASA Books program. What a wonderful treatise it is, filled with theory, data, explanations, and references. I begin each acoustics course by telling the students that they will hate using his book in class, because it is more intense than most books they will encounter, but they will recognize upon completion of their studies that it has made them an acoustics expert.

Perhaps education will be his greatest legacy, and yet there is much more that is important. Allan is tireless in his professional activities, especially within our Society. Is there an area where he has not participated? Is there an activity where he has not taken a leadership role? Technical committees, Executive Council, Education, Membership, Chair of the American Society of Mechanical Engineers (ASME) Noise Control and Acoustics Committee, US National Committee for Theoretical and Applied Mechanics, and countless other activities that do not show up on a resumé. He was an area editor for our *Journal*, founding co-editor of the *Journal of Computational Acoustics* and co-editor of the Physical Acoustics series published by Academic Press. Presently he is a member of the Editorial Board of the *Journal of Sound and Vibration*. All this while serving as Editor-in-Chief for our *Journal*. He has taken over in a perilous time, with technology threatening our old ways. He has launched major initiatives to ensure that our *Journal* continues to be the premiere vessel for disseminating knowledge of acoustics. He has converted a part-time job into a full-time avocation. There are no better hands than his to guide the *Journal* into the future.

If Allen has an idol, it is Lord Rayleigh, which is a demonstration of the standards that he sets for himself and others. Allan delivered the Rayleigh Lecture at the ASME annual meeting in 1992. In November of last year at our meeting in San Diego, he was the first recipient of the Rossing Prize in Acoustics Education, for which he gave an enthralling talk on the life of Rayleigh. An incident when he was on the faculty of Georgia Tech vividly illustrates Allan's esteem for Rayleigh. We were conducting the oral part of the Ph.D. qualifying exam in acoustics. One student had done quite well, but we had some time left, so Allan asked him "Who was the greatest acoustician?" The student had studied everything, but not in a historical context, and clearly was flustered. After thinking for approximately a minute, he said to Allan "You!" Allan corrected him by simply saying "Rayleigh", then thanked him for the compliment. The student passed despite this serious gap in his knowledge!

Other recognitions Allan Pierce has received include the Senior U.S. Scientist Award from the Alexander von Humboldt Foundation in 1976, the Silver Medal in Physical Acoustics from ASA in 1991, and the Per Bruel Gold Medal for Noise Control and Acoustics from ASME in 1995. He was named the Outstanding Thesis Advisor by the Georgia Tech Chapter of Sigma Xi in 1988. I do not think that any one now alive has done as much as he to advance acoustics, not just as an area in which to work, but as a place in which to make a home. It is most appropriate that the prior recognitions of Allan Pierce's contributions be capped by our Gold Medal.

JERRY H. GINSBERG

Session 4aAA**Architectural Acoustics and Noise: Preserving Acoustical Integrity in the Course of Renovation**

Daniel R. Raichel, Chair
 2727 Moore Ln., Fort Collins, CO 80526

Chair's Introduction—8:30*Invited Papers***8:35**

4aAA1. Renovating Teatro alla Scala Milano for the 21st century, Part I. Higiní Arau (Travesera de Dalt, 118, 08024 Barcelona, Spain, h.arau@arauacustica.com)

Teatro alla Scala of Milan, known simply as La Scala throughout the world, is an old but venerable opera house that achieved legendary status in the world of music. A great number of singers of Olympian status have sung there; and major operas, among them Verdi's Falstaff and Otello and Puccini's Turandot, premiered there. The 227-year-old theater is beloved with a passion by the Milanese and the Italians, but it has suffered the ravages of time. It needed to be renovated in order to reverse material decay, meet current fire codes and security requirements, incorporate a new HVAC system, and to accommodate badly needed modern stage machinery. This renovation project took 3 years during which the theater was closed, and it included the construction of an elliptical 17-floor fly tower, designed by architect Mario Botta, for housing rehearsal rooms and serving as a scenery changing facility. The renovation proposal originally aroused a strong sense of melodrama among the extremely excitable Italian opera buffs who feared the desecration of their beloved edifice, but the acoustics and the beauty (carried out by Elisabetta Fabbri Architect) of the auditorium were preserved (and even enhanced). In this paper we explain how this project was achieved.

8:55

4aAA2. Renovating Teatro alla Scala Milano for the 21st century, Part II. Higiní Arau (Travesera de Dalt, 118, 08024 Barcelona, Spain, h.arau@arauacustica.com)

The acoustic phase of La Scala renovation began in September 2002, after the main stalls and other sections of the theater were demolished. This assignment was twofold: (a) design of the auxiliary building with architect Mario Botta, and more importantly, (b) collaboration with architect Elisabetta Fabbri in restoration of the auditorium through acoustic analyses of proposed solutions. Only one set of acoustical measurements was known to be taken before demolition; and reliance had to be placed on hearsay from audience members. The author used his own computer program that included some of the salient features of other programs such as Odeon, Epidaure, Raynoise, etc. but avoided their pitfalls. This program was the only one that correctly predicted the known RT of the auditorium through the use of H. Arau Purchades formula [Arau, H., 1988. *Acustica*. Hirzel Verlag **65**(4), 163–180] and the authors dimension theory [Arau, H. 1997. Variation of the reverberation time of places of public assembly. *Building Acoustics* **4**(2)]. A new floor was designed to provide sufficient vibration transmission to the audience, actuating as a radiation box installed to direct sound vertically. Music Director Ricardo Muti pronounced the acoustical results as being excellent.

9:15

4aAA3. Maintaining the acoustics of Boston Symphony Hall. Robert Berens, Benjamin Markham, and Carl Rosenberg (Acentech Inc., 33 Moulton St., Cambridge, MA 02138, rberens@acentech.com)

Boston Symphony Hall celebrated its centennial in 2000; one of its biggest birthday presents was a new organ. More correctly, the partially new and partially refurbished organ made its debut in October 2004, for James Levine's inaugural concert as Music Director of the Boston Symphony Orchestra. Symphony Hall's original organ was replaced in 1949, but its replacement was never regarded as a great concert instrument. While some wanted the new organ work to include turning the Hall into an organ showcase, the BSO management was adamant that the acoustics of the Hall not be affected. Acentech was asked to provide technical oversight to help minimize any such changes. Tests were conducted to assess the potential impacts that removal and renovation of the organ pipes could have on Symphony Hall's acoustics. Reverberation times were measured in the organ chamber, on-stage, and in the Hall. Remedial measures were devised for use, if necessary, to counteract the expected changes in the organ chamber acoustics as the pipes were removed. However, such changes showed little effect on the acoustics of the stage or the hall. This paper describes the work done during the organ renovation process to protect the hallowed acoustics of Symphony Hall.

9:35

4aAA4. Historical preservation of acoustics at Spelman College, Atlanta. Carl Rosenberg, Benjamin Markham (Acentech Inc., 33 Moulton St., Cambridge, MA 02138, crosenberg@acentech.com), and Barbara Kovacs Black (Surber Barber Choate & Hertlein Architects, Inc., Atlanta, GA 30309)

In 1926, Spelman College in Atlanta dedicated its new Sisters Chapel, named after Laura Spelman Rockefeller and Lucy Maria Spelman, mother and aunt of John D. Rockefeller, Jr., the major donor. The Chapel has around 1000 seats and serves as a major venue for a myriad of programs, ceremonies, concerts, and student activities. In 2003, architects Surber Barber Choate and Hertlein were commissioned to renovate and rehabilitate the chapel with air-conditioning, new toilet facilities, new roof, enlarged balcony, and new sound system. The neo-classic pseudo Georgian basilica has a barrel vault running the length of the room. Sometime in the past, probably around 1950, the chapel ceiling was covered with a sound absorbing acoustic tile, presumably to rectify a focusing problem from the curved ceiling. With computer modeling, Acentech reconstructed the apparent acoustical problems and used this information to guide a renovation. The goal was to recapture the visual aspects of the original design, enhance the reverberation for the renovated Holtkamp organ and singing, and at the same time avoid any focusing problems for which the acoustic tile was presumably added. The project included mechanical system noise and vibration control and a new highly directional sound system.

9:55

4aAA5. Barnum Hall—The continuing renovation of a Streamline Moderne theater. Neil A. Shaw (Menlo Sci. Acoust., Inc. P.O. Box 1610, Topanga, CA 90290-1610, menlo@ieee.org), Kenneth Koslow (Santa Monica Malibu Unified School District, Santa Monica, CA 90404), Jean Sedillos (Restore Barnum Hall, Santa Monica, CA 90403), and Jim Mobley (Renkus-Heinz, Foothill Ranch, CA 92610)

Barnum Hall, built by the WPA in 1937 and located on the campus of Santa Monica High School, was the first Santa Monica Civic Auditorium. After 60 years it was closed for renovations. The building has historic significance, which placed some limitation on the renovation design as to what could be done. Balancing the interests of the various stakeholders—the Santa Monica-Malibu Unified School District, the City of Santa Monica Landmarks Commission, fund raisers, high school faculty, and community members, among others—impacted the design and construction process. The first phase included expanding and modernizing the stage house and surrounding support areas. Phase II of the work was concerned with mitigating the orchestra level rear wall reflections, modifying the balcony nose profile, adding an orchestra pit lift, the purchase of a custom acoustical orchestra shell, and improving creature comfort (air conditioning) and amenities (improving restrooms, creating a backstage area, and refurbishing 75% of the theater seats). A study of several sidewall shaping schemes were then analyzed to determine if the improvement in the level of envelopmental sound warranted the cost. Phase III will see the installation of a modern sound reinforcement system and refurbishing the last section seats.

10:15–10:30 Break

10:30

4aAA6. Renovating cultural icons. Gregory Miller (Talaske, 105 N Oak Park Ave., Oak Park, IL 60301, greg@talaske.com)

Three case studies of historic renovations are presented where acoustics were a key component of the renovation process. Each hall is an icon of the cultural life in the surrounding community. The first case study is Troy Savings Bank Music Hall, a late nineteenth century musical gem where recent renovations (adding variable acoustic features) required that the unamplified acoustics be unaltered. The second example, the Coronado Theatre in Rockford, Illinois, illustrates the sensitive modification of the acoustics in a historic vaudeville house adapted to modern multi-use requirements. Finally, The Great Hall at The Cooper Union in New York City, where Abraham Lincoln delivered his great “Right Makes Might” speech in 1860, is presented as an example of a renovation that utterly destroyed a historic acoustic environment, and discusses how this can be avoided.

Contributed Papers

10:50

4aAA7. Returning an acoustic legend to its original glory (and then some). Scott Pfeiffer (Kirkegaard Assoc., 801 W. Adams St., 8th Fl., Chicago, IL 60607)

Hill Auditorium at the University of Michigan broke all of the rules of modern day acoustic design, while still providing a tremendously rewarding experience for large audiences. The combination of diverse and intriguing programming by the University Music Society and the University Music Department, including the use of the Frieze Memorial Organ, provides tremendous opportunities for the community of Ann Arbor, MI to access great performances. The restoration project included air conditioning, elevators, and new sound systems. Additionally, improvements were made with respect to seating; specifically, sightlines, accessible seating, and under balcony seating were improved upon. This restoration project provided opportunities to revitalize the already famous acoustics without sacrificing unique and characteristic qualities of the hall. By drawing on

the original architects, Albert Kahn Associates, utilizing the expertise of Quinn Evans Architects, as well as through the help of Gary Steffy Lighting, Fisher Dachs Associates, and many others in design and construction, (most notably a great client in the University of Michigan), the project secured the 2004 AIA National Honor Award.

11:05

4aAA8. Preserving the acoustics of the Mahaiwe Theater. Ronald Eligator (Acoust. Dimensions, 145 Huguenot St., New Rochelle, NY 10801, religator@acousticdimensions.com)

The Mahaiwe Theater, Great Barrington, Massachusetts, opened in 1905. It has been home to vaudeville, traveling Broadway shows, opera, concerts, and movies. In 2003, the theater began a multi-phase renovation project. The renovation project includes restoration of historic finishes,

4a THU. AM

new seating, installation of air conditioning and heating systems, and installation of sound and projection systems. A stage shell for music performances will also be installed. A key aspect of the renovation is the preservation of what has been widely regarded as excellent acoustics. Some aspects of the original structure contributed to this reputation (e.g., room shaping, finishes), while others undoubtedly were less helpful (e.g., the 1905-era ventilation system). The renovation provides an opportunity to preserve the positive attributes of the room, while correcting the deficiencies. Among the key issues discussed are the design of historically-sensitive finishes that respect the acoustic requirements of the multi-use space, acoustical design of a new quiet air conditioning system, and design of an enclosure for music ensembles.

11:20

4aAA9. Preservation of the acoustics of the Salt Lake Tabernacle, Lessons learned from acoustical characterization of the hall. Sarah Rollins and Timothy Leishman (Acoust. Res. Group, N283 ESC, Brigham Young Univ., Provo, UT 84606, sr223@email.byu.edu)

Current and historical properties of the Salt Lake Tabernacle have been investigated to characterize its unique acoustics. This characterization will be used to help maintain the current acoustical conditions after a seismic renovation. This paper discusses computer models developed for this purpose and impulse response measurements taken to quantify the acoustics in their current state. It also explores spatial variation of room acoustic parameters derived from the models and measurements.

THURSDAY MORNING, 19 MAY 2005

PLAZA C, 8:15 TO 11:50 A.M.

Session 4aAB

Animal Bioacoustics: Tools for Animal Bioacoustics: New Designs and Directions I

William C. Burgess, Chair
Greeneridge Sciences, 6060 Graham Hill Rd., Felton, CA 95018

Chair's Introduction—8:15

Invited Papers

8:20

4aAB1. A binaural acoustic recording tag reveals details of deep foraging in beaked whales. Mark Johnson, Peter Madsen, Peter Tyack (Woods Hole Oceanograph. Inst., Woods Hole, MA 02543, majohnson@whoi.edu), and Natacha Aguilar de Soto (Univ. La Laguna, Tenerife, Spain)

A new acoustic and orientation recording tag (DTAG) for marine mammals contains two hydrophones, 25 mm apart, sampled at 192 kHz. Data are stored with loss-less compression in 6.6 GB of memory giving a 9.5 h recording capacity. Stereo DTAGs have been attached with suction cups to Cuvier's and Blainville's beaked whales in Liguria and the Canary Islands. Both species make regular clicks centered at 40 kHz during foraging dives. The DTAGs recorded clicks from the tagged whale and other whales nearby as well as echoes from targets in the water. The angle-of-arrival (AoA) of these sounds was determined both by cross-correlation and time-delay fitting of signals from the two hydrophones on the tag. The AoAs of clicks from tagged whales reveal that they turn their heads from side-to-side while foraging while the AoAs of echoes are consistent with a beam-width of about 20°. By scanning a narrow acoustic beam, the whales may be able to search a large water volume while reducing clutter from multiple echoes. Frequent clicks from untagged whales with distinct AoAs indicate that several whales forage together during deep dives and such group cohesion may be a contributing factor to strandings of these species related to use of naval sonar.

8:40

4aAB2. Miniature self-contained acoustic recorders applied in a survey of beluga-whale populations in Knik Arm, Alaska. William C. Burgess (Greeneridge Sci., 6060 Graham Hill Rd Ste F, Felton, CA 95018), Michael T. Williams (LGL Alaska Res. Assoc., Anchorage, AK 99518), and Susanna B. Blackwell (Greeneridge Sci., Aptos, CA 95003)

Since 1995, acoustic recording tags attached to marine wildlife have increased our understanding of how animals use and respond to sound. Tagging studies require these instruments to be designed for minimum size and maximum recording capability; however, by designing for ease of use as well, a variety of other acoustic studies become feasible that could not otherwise be attempted for lack of time, staff, or support. As a result, designing for ease of use multiplies the scientific power of acoustic recording tag technology. In late 2004 four acoustic recording tags, known as Bioacoustic Probes and developed by Greeneridge with ONR support, were used as fixed recorders in a primarily visual survey of beluga whales in the Knik Arm of Cook Inlet, Alaska. The study took place with a compressed schedule and was staffed by field biologists with little time available for technical training on the use and maintenance of the instruments. Designing the instruments for ease-of-use enabled the study to include an independent acoustic component for assessing the presence or absence of beluga whales in areas that were being surveyed visually during daylight hours. [Work supported by Knik Arm Bridge and Toll Authority.]

9:00

4aAB3. Modular autonomous array deployments using passive acoustic time synchronization. Aaron Thode (Marine Physical Lab., Scripps Inst. of Oceanogr., 9500 Gilman Dr., MC 0238, La Jolla, CA 92093-0238, thode@mpl.ucsd.edu) and William Burgess (Greeneridge Sci., Felton, CA 95018)

Multi-element acoustic arrays generally use wired cable to transmit signals to a central recording location. While convenient, hardwiring hydrophones together increases array fragility and field costs while decreasing deployment flexibility. It is thus difficult to integrate acoustic array operations with field activities involving marine mammals in remote environments. Economic trends in the cell-phone and consumer electronics industries, combined with recent trends in tagging technology development, have led to the existence of compact low-power autonomous acoustic recorders that store data to either flash memory or small hard drives. In this presentation it is shown how two or more autonomous recorders can be time-synchronized using passive acoustic measurements of the background ambient noise field, effectively creating coherent array processing systems of varying aperture, spacing, and deployment geometry. Configurations tested to date include short and large-aperture vertical arrays off the coasts of Australia and Alaska, bottom-mounted horizontal arrays in San Ignacio Lagoon, Baja California, and instruments installed inside a glider. [Work sponsored by ONR Acoustic Entry Level Faculty Award.]

9:20

4aAB4. High-frequency Acoustic Recording Package (HARP) for long-term monitoring of marine mammals. Sean Wiggins, Chris Garsha, Kevin Hardy, and John Hildebrand (Scripps Inst. of Oceanogr., 9500 Gilman Dr. MPL-0205, La Jolla, CA 92093-0205)

Advancements in low-power and high-data-capacity computer technology during the past decade have been adapted to autonomously record sounds from whales over long time periods. Acoustic monitoring of whales has advantages over traditional visual surveys including greater detection ranges, continuous long-term monitoring in remote locations under various weather conditions, and lower cost. One currently used tool for providing long-term acoustic monitoring of marine mammals is an autonomous acoustic recording package (ARP) which uses a tethered hydrophone above a seafloor-mounted instrument frame. Since 2000, ARPs have been deployed to record baleen whale sounds in the Bering Sea, in the Beaufort Sea, in the Gulf of Alaska, off the coast of southern California, around Antarctica and near Hawaii. ARP data have provided new information on the seasonal presence, abundance, call character and patterns of calling whales. The need for a broader-band, higher-data capacity system capable of recording odontocete whales, dolphins and porpoises for long time periods has prompted the development of a High-frequency Acoustic Recording Package (HARP). The HARP design is described and data analysis strategies are discussed using examples of HARP broad-band (sample rates up to 200 kHz) recorded data.

9:40

4aAB5. A near-real-time acoustic detection and reporting system for endangered species in critical habitats. Christopher W. Clark, Thomas Calupca (Bioacoustics Res. Program, Cornell Lab. of Ornithology, 159 Sapsucker Woods Rd., Ithaca, NY 14850), Douglas Gillespie (Intl. Fund for Animal Welfare, Yarmouth Port, MA 02675), Keith Von der Heydt, and John Kemp (Woods Hole Oceanograph. Inst., Woods Hole, MA 02543)

Passive acoustics is an effective mechanism for detection and recognition of species-specific sounds and can be a more cost-effective approach than visual techniques for monitoring populations of rare or endangered species. A network of moored buoys has been strategically deployed in and around Cape Cod Bay to report detections of northern right whales in critical habitat. Each buoy continuously and automatically monitors for right whale contact calls and transmits detection and ambient noise data by cell or satellite phone to Cornell University on a regular basis. Each day, validated data are automatically unloaded into a Website database to provide on-line graphical and numerical data summaries. The array of three buoys deployed in the Bay will eventually be synchronized to allow localization and tracking of individual animals. [Work supported by funds from the NOAA Right Whale Grants Program and augmented by funds from the Commonwealth of Massachusetts Division of Marine Fisheries.]

10:00–10:10 Break

10:10

4aAB6. Extensible bioacoustical analysis software: Two examples. Harold G. Mills and Harold K. Figueroa (Bioacoustics Res. Program, Cornell Lab. of Ornithology, 159 Sapsucker Woods Rd., Ithaca, NY 14850, hgm1@cornell.edu)

Two extensible software tools for bioacoustical analysis have been developed by the Bioacoustics Research Program at the Cornell Laboratory of Ornithology. Each tool provides a framework for sound visualization and analysis, supporting both manual selection and automatic detection of acoustic events in arbitrarily long recordings, manual annotation and automatic measurement of properties of these events, and management of the created acoustic metadata. The acoustic events of interest might include animal vocalizations and/or anthropogenic sounds. Simple examples of measurable event properties include duration, total energy, and spectrum. More advanced measurements might classify sounds according to the species or individuals that produced them, or estimate the location of a vocalizing animal. Both software tools can be extended in various ways, including the addition of new detectors and measurements. It has been found that this extensibility adds great value to the tools by allowing them to be readily adapted for new applications by bioacousticians or others with programming ability. It has also been found that such tools greatly facilitate the development of new detection and measurement algorithms by allowing prototype algorithms to be rapidly evaluated. Several case studies are presented that illustrate these principles. [Work supported in part by NSF.]

4a THU. AM

4aAB7. Measurement of biosonar signals of echolocating bat during flight by a telemetry system. Hiroshi Riquimaroux (Dept. of Knowledge Eng. and Comput. Sci., Doshisha Univ., Kyotanabe, Kyoto 610-0321, Japan)

A telemetry microphone system (Telemike) has been developed, which can be mounted on the bat's head, to measure acoustic characteristics of emitted pulses and returning echoes. The system allows us to monitor what the bat listens to during its flight. A high-speed video camera system has been also adopted together with the Telemike to trace positions in space of the flying bat for analyzing temporal pulse emission patterns. With those devices, how the CF-FM bats execute parallel time-sharing real-time processing during their flight can be investigated. Doppler-shift compensation, echo amplitude compensation and processing for multiple target detections were observed. Some of the evidence found by the Telemike will be introduced and discussed. [Work partly supported by the Innovative Cluster Creation Project promoted by MEXT and by a grant to RCAST at Doshisha University from MEXT.]

Contributed Papers

10:50

4aAB8. 3D modeling and animation techniques to elucidate mechanisms of echolocation-based acoustic perception. Seth Horowitz (Univ. at Stony Brook, Dept Psychiatry, HSC T-10 Rm. 086, Stony Brook, NY 11794, shorowitz@neuropop.com) and James Simmons (Brown Univ., Providence, RI 02912)

Current models of echolocation focus on the presence of specular acoustic surface reflections or "glints" as the perceptual cues that provide information about object distance and geometry. Using current material and motion 3D modeling techniques for visual animation, we have developed models of common real-world elements that can be detected by bat biosonar. By mapping acoustic reflectivity data from natural and artificial targets onto luminosity, specularity and reflection coefficients for materials and accurate spatial models of these objects, we have generated visual glint-like reflections from object surfaces that are similar to echolocation data. This allows us to create 3D visual animations of echolocation auditory scenes. Using these animations, human observers can integrate object shapes, sizes and movements based on gestalt grouping, particularly when the target moves or the echolocation source is in moving in relation to the target. These types of modeling tools may help elucidate the perceptual phenomena arising from integration of the individual acoustic glint structures that allow bats and other echolocating animals to create complex umwelts from auditory data.

11:05

4aAB9. The acoustic environment of the southern resident killer whales in Haro Strait: Propagation modeling and analysis of field measurements. Christopher Jones and Michael Wolfson (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

We will discuss the complicated shallow water acoustic environments of the southern resident killer whales in the Haro Strait of the Puget Sound through combined analysis of field measurements and acoustic propagation modeling in the frequency range of 1 to 20 kHz. Haro Strait is a highly variable acoustic environment with active commercial shipping, whale watching, and Naval activity. Southern resident killer whales are of unique public concern in this area because of potentially high impact by anthropogenic sound sources, either by auditory masking that may interfere with foraging strategies or annoyance and disorientation due to complicated reverberation. Predictive acoustic modeling in combination with field measurements can aid in understanding the mechanisms of impact and better inform assessment of the risk, providing a critical step towards the quantitative evaluation of impact in the context of complicated environments, changing background sound levels, and emerging management issues. Preliminary results of modeling and analysis of data collected in the summer of 2004 will be presented.

11:20

4aAB10. Ship strike acoustics. Edmund Gerstein, Joseph Blue (Leviathan Legacy Inc., 1318 SW 14th St., Boca Raton, FL 33486), and Steve Forsythe (Naval Undersea Warfare Ctr., Newport, RI)

The confluence of Lloyds mirror effect together with acoustical shadowing and spherical spreading pose significant detection challenges for whales and manatees. Direct measurements of approaching vessels using horizontal and vertical hydrophone arrays demonstrate how the noise from vessels can become indistinguishable from ambient noise. Geometric scattering is also estimated for various hull dimensions as the acoustical shadows cast ahead of ships provide no warning to animals near the surface within the shadow zone. Animals at sufficient depths, or outside the shadow boundaries that hear approaching vessels, may seek refuge near the surface or directly in the path of vessels where it is relatively quiet. Speed reductions proposed to reduce collisions do not address the underlying acoustical challenges marine mammals face. Field measurements support predictions that noise intensity is proportional with vessel speed to the 5th power. In multiple ship environments the acoustical masking challenges are greatest and noise from slow vessels operating in marine mammal corridors can mask the sounds of distant faster moving vessels. A low intensity bow-mounted projection system has been designed to selectively fill-in acoustical shadows with modulated ship noise to mitigate masking and near surface effects, and neutralize the dangerous ambiguity posed by acoustical shadows.

11:35

4aAB11. Detection ranges for acoustic based manatee avoidance technology. Richard Phillips (Dept. of Mech. and Aerosp. Eng., Univ. of Florida, Gainesville, FL 32611-6250), Christopher Nierecki (Univ. of Massachusetts Lowell, Lowell, MA 01854-2881, Chris_Nierecki@uml.edu), and Diedrich O. Beusse (Univ. of Florida, Gainesville, FL 32610-0126)

The West Indian manatee (*Trichechus manatus latirostris*) has become endangered partly because of watercraft collisions in Florida's coastal waterways. Several boater warning systems, based upon manatee vocalizations, have been proposed to reduce the number of collisions. One aspect of the feasibility of an acoustically based system relies upon the distance at which a manatee vocalization is detectable. The magnitude of environmental noise and manatee vocalizations, as well as the acoustic spreading properties of the habitat will help to estimate the detection range of a vocalizing manatee. This study combines measured source levels of manatee vocalizations with the modeled acoustic properties of manatee habitats to develop a method for determining hydrophone spacing requirements for acoustically reliant manatee avoidance technologies. In quiet environments (background noise 70 dB) it was estimated that manatee vocalizations are detectable at approximately 250 m, with a 6 dB detection threshold. In louder environments (background noise 100 dB) the detection range drops to 2.5 m. Noise generated by boating traffic is also investigated. In a habitat with 90 dB of background noise, a passing boat with a maximum noise floor of 120 dB would be the limiting factor when it was within approximately 100 m of the hydrophone.

Session 4aNS**Noise, ASA Committee on Standards and Engineering Acoustics: Noise Control Feasibility: Technical, Legal and Economic Issues**

John P. Seiler, Cochair

Mine Safety and Health Administration, Cochran Mill Rd., Pittsburgh, PA 15236

Angelo J. Campanella, Cochair

*Campanella Associates, 3201 Ridgewood Dr., Hilliard, OH 43026-2453***Chair's Introduction—9:00*****Invited Paper*****9:05****4aNS1. Industrial noise control—Technical and economic feasibility.** Stephen I. Roth (Roth Acoust. Assoc., 2352 Norton Rd., Pittsburgh, PA 15241, sroth@rothacoustics.com)

Industrial noise control is too often considered technically not possible or economically infeasible. The use of ear protection devices becomes the fallback approach to attempt to reduce the risk of worker hearing loss. In reality there are noise control opportunities in industrial settings than are low cost and provide significant noise reduction without affecting production and maintenance practices. This paper will present mechanical and pneumatic noise controls that can be successfully applied to industrial environments. This paper will also present information on OSHAs approach to determine whether engineering noise controls are considered economically feasible.

Contributed Papers**9:25****4aNS2. The feasibility of noise and reverberation control for good acoustics in classrooms.** Bennett M. Brooks (Brooks Acoust. Corp., 27 Hartford Turnpike, Vernon, CT 06066, bbrooks@brooksaoustics.com)

The acoustics of classrooms has been the subject of increasing scrutiny, as good speech intelligibility is recognized as an essential feature of an effective educational environment. There is consensus on the need for low levels of background noise and reverberation. Acceptable values for these parameters have been published as a standard [ANSI S12.60-2002]. The technical feasibility of achieving these goals is well established. In the US, few jurisdictions require that classrooms meet acoustical standards. The limits to constructing acoustically adequate classrooms are primarily economic. Perception and understanding of the issue are also important, as those who specify and pay for classrooms, school boards and taxpayers, are not those directly served in them, students. As the direct customers and their immediate providers, teachers, have little power, feasibility depends on competing priorities set by interested non-users. Therefore, advocates must sell the cost/benefits to the decision makers. Widely scattered data are available on costs as a percentage of the total school building project. However, detailed cost data for individual mechanical system components that meet noise standards are scarce. Also, long-term economic paybacks

must be quantified. Available cost data and areas needing better definition are summarized here.

9:40**4aNS3. Low impact renovations with high performance sound insulation.** Kenneth P. Roy and Sean D. Browne (Armstrong Innovation Ctr., 2500 Columbia Ave., Lancaster, PA 17604)

Architectural renovations, although often driven by visual objectives, are a good opportunity to pursue acoustical interventions involving both sound intrusions into a space, and sound quality of the acoustic environment within the space. A high priority is to keep the construction impact low as it affects both project cost and schedule. Obviously, if the architectural intervention can be accomplished using external measures such as the addition of added elements to the existing structure, then that would be a definite advantage. Test data will be presented for the application of sound insulating/absorbing panels to existing drywall structures. It will be shown that such an overlay panel can add sufficient damping to result in essentially "mass law" performance of the system. This will not only result in STC improvement, but more importantly will result in even higher voice insulation between spaces. This effect will be auralized for comparison listening purposes. If the added panel is also of a sound absorbing nature, such as those used in the tested cases, then they will also have the effect of adding to the receiving room sound performance.

9:55

4aNS4. Structural vibration isolation design for a magnetic resonance imaging (MRI) system. Chad Himmel (JEAcoust., 1705 W Koenig Ln, Austin, TX 78756, info@jeacoustics.com)

This structural vibration control case study presents problems, constraints and design solutions for existing magnetic resonance imaging system (MRI) installation in a medical office facility. Objective: Reduce MRI noise received in adjacent, unrelated spaces. Manufacturer's data indicated that airborne MRI sound emissions could exceed permissible noise criteria for nearby occupied rooms. Experience with MRI systems indicated possibility of structure borne vibration that could result in radiated noise in other spaces. Structure borne vibration paths needed attenuation or isolation, to prevent excessive or annoying and distracting noise to adjacent office spaces. In addition, containment design was required for loud noises in the magnet room. Measurements revealed transient vibration of building floors, walls, and ceilings coincident with MRI operation. Measurements also revealed transient airborne noise levels consistent with radiated sound produced by vibration acceleration of large surfaces. Spectral analyses of noise and vibration led to design of a retrofit vibration isolation mounting scheme for the MRI on existing building floor. Design parameters included structural resonance of slab, manufacturer's allowable vibration for MRI, and receiver room permissible criteria. Noise and vibration measurements will show building conditions before and after implementation of retrofit vibration isolation mounts. Plan and section drawings will illustrate design solutions.

10:10–10:30 Break

10:30

4aNS5. MRI system vibration and noise considerations in hospital design. Eric E. Ungar and Jeffrey A. Zapfe (Acentech, Inc., 33 Moulton St., Cambridge, MA 02138, eungar@acentech.com)

Magnetic Resonance Imaging (MRI) systems are increasingly being located on above-grade floors of hospital buildings. The floor structures that support these systems need to provide vibration environments that satisfy the system's criteria, and airborne and structure-borne transmission of noise to neighboring areas needs to be taken into account in order to provide acceptable acoustical environments in these spaces. The fundamental operating principles of MRI systems are reviewed in relation to the systems' susceptibility to vibration and to their sound producing mechanisms. Vibration criteria for some widely used systems are discussed and spectra of airborne and structure-borne noise are illustrated. Means for vibration control and for the attenuation of structurally transmitted sound are delineated, together with architectural considerations for limiting noise transmission to spaces adjacent to MRI suites.

10:45

4aNS6. Psychoacoustical comparison of active versus passive noise control techniques. Gerard Mangiante and Georges Canevet (Laboratoire de Mécanique et d'Acoustique, CNRS, 31 Chemin Joseph-Aiguier, 13402 Marseille Cedex 20, France)

Comparisons of active versus passive noise control techniques can be found in various papers. However, this comparison mainly concerns economic, operational, or technical considerations. The present contribution aims at describing the psychoacoustical effects produced by some of the classical solutions used in passive and active noise control. The models introduced by Zwicker and his coworkers, and by Moore and Glasberg are used to evaluate the auditory efficiency of passive and active noise control techniques. Several types of signals were examined: (i) test signals obtained with a band of noise embedded in a white noise or a pink noise; (ii) actual environmental noises: noise produced by the turbine of an aircraft or by a car engine, and several interior noises (locomotive, helicopter and car). It is shown that the modifications in the spectrum of a signal that can be produced by active control are sometimes disappointing, because they

induce a subjective enhancement of the high-frequency portion of the spectrum. What the listener then commonly reports is that overall the signal has become slightly softer, but also more unpleasant. The use of a hybrid noise control technique, combining active and passive control, can greatly reduce this effect.

11:00

4aNS7. Adaptive multi-modal active noise control. Adam K. Smith, Jeffrey S. Vipperman, and Daniel D. Budny (Univ. of Pittsburgh, Dept. of Mech. Eng., 648 Benedum Hall, Pittsburgh, PA 15261)

Low frequency Active Noise Control (ANC) has been found to work well for sound suppression in acoustic ducts. Feedforward and feedback are the two distinct methods in which to implement (ANC) schemes. A feedforward implementation is realized by utilizing a known disturbance measurement (e.g. with a microphone), and using the measurement to compute the actuator (loudspeaker) signal. Feedback control modifies dynamics of an enclosed sound field to add damping. For example, one can use a resonant filter in the feedback path that is tuned to a selected mode in the enclosed sound field to be controlled. The resonant filter model is similar to those used for positive position feedback (PPF) in structures. A method of compensating the strong speaker dynamics (phase angle) was recently investigated by Bisnette and Vipperman (2004). The method was found to improve the performance and stability of the controller. Here, the above mentioned procedure is expounded upon to provide improved multi-modal performance, using higher order band-pass filters. Methods for adapting the controller parameters (e.g., gain, frequency, bandwidth) are also presented.

11:15

4aNS8. Acoustic power and intensity control by absorptive materials arrangement in an enclosure. Sung-Ho Cho and Yang-Hann Kim (Dept. of Mech. Eng. KAIST, 373-1 Sci. Town, Daejeon-shi, Korea)

This paper studies how the sound distribution of room or cavity is affected by the absorptive materials arrangement on the wall. In other words, we want to know, very specifically, how the change of boundary condition affects rooms sound distribution. The boundary condition can be modified by changing not only absorbers position but also its acoustic impedance. The effect of changing boundary condition is expressed in terms of modal admittance on the enclosures surface. What we have gotten from this study is that the absorptive materials placement is closely related to wave length, cavity size and geometry. The authors pay attention to the intensity field, especially reactive intensity divergence in space. The divergence of reactive component of intensity is directly related to acoustic energy distribution which is composed of potential and kinetic energy. Using acoustic energy balance equation, the relation between global noise control performance and absorptive material arrangement is explained. In this point of view, the possibility of global noise reduction will be discussed in terms of acoustic potential energy reduction.

11:30

4aNS9. Sound attenuation by lattices of rigid elliptic cylinders. Daniel Torrent, Andreas Håkansson, and José Sánchez-Dehesa (Nanophotonics Technol. Ctr., Tech. Univ. of Valencia, E-46022 Valencia, Spain, jsdehesa@upvnet.upv.es)

The attenuation properties of acoustic barriers consisting of a few layers of rods with elliptical cross section has been studied by multiple scattering theory. Their performance as a function of number of layers and orientation of the ellipses will be reported and compared with similar structures based on cylinders of circular section. A comparison with available experiments will also be presented. Finally, a genetic algorithm is applied to optimize the attenuation in a wide range of frequencies. [Work supported by MEC of Spain.]

Session 4aPAa**Physical Acoustics, Biomedical Ultrasound/Bioresponse to Vibration and ASA Committee on Standards: Cavitation and Other Mechanical Effects in Biomedical Ultrasound: A Special Session to Honor the Work of Wesley Nyborg I**

Lawrence A. Crum, Cochair

Applied Physics Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698

Junru Wu, Cochair

*Dept. of Physics, Univ. of Vermont, Burlington, VT 05405***Chair's Introduction—7:55*****Invited Papers*****8:00****4aPAa1. Wes Nyborg: Scientist, role model, and friend.** Marvin C. Ziskin (Ctr. for Biomed. Phys., Temple Univ. Med. School, 3420 N. Broad St., Philadelphia, PA 19140)

Wesley L. Nyborg obtained his Ph.D. degree in physics at Pennsylvania State College in 1947. He has held faculty positions at Penn State, Brown University, and since 1960 has been at the University of Vermont where he is presently Emeritus Professor. Dr. Nyborg's entire career has been devoted to biophysical acoustics: low frequencies at first, but since the early 50s ultrasonic frequencies of biomedical interest. He has made significant contributions in many areas of ultrasound biophysics, especially in cavitation and other non-thermal mechanisms relevant to biological effects. He has also derived a number of important relationships that are widely used today in determining temperature elevation in clinical ultrasound examinations. He has been the role model and friend to many colleagues, and has received many awards.

8:20**4aPAa2. Sound sources and propagation.** Mahlon Burkhard (157 Stephen Ln., Charles Town, WV 25414)

From July 1945 through May 1950, the Acoustics Laboratory of the Department of Physics at Pennsylvania State [College] University conducted "investigations concerning the production and propagation of sound, both sonic and ultrasonic, in the lower atmosphere, through the ground, and in other media." Wesley L. Nyborg was a key member of the group of 53 researchers employed in the program at some-time during the period. A sound source used extensively in the study of ultrasonic propagation in the atmosphere was a small whistle. "Edge tones" generated by a jet of air impinging on a narrow object are phenomena critical to the operation of these whistles as well as organ pipes and many sounds associated with fluid flow. Because the sound can be controlled and levels can be quite high, one proposed application was as a source for radiating small quantities of liquids with ultrasound. In addition to developing a physical understanding of this sound generation mechanism, Nyborg participated in studies and analysis of the influence of micrometeorology on propagation of high frequency sound in the atmosphere.

8:40**4aPAa3. Infrasonic tonal resonances of tropical hurricanes.** Samuel A. Elder (Phys. Dept., U.S. Naval Acad., Annapolis, MD 21402)

In news pictures saturating the media last summer there is contained subtle but compelling evidence that tropical hurricanes emit intense and extremely low-frequency tones, originating in the eye. The excellent symmetry of a well-formed eye predisposes the central cavity of a hurricane to resonate in the normal modes of a low- Q nearly cylindrical chamber, due to the slight mismatch in sound speed at the eye wall. Second order perturbation analysis of interaction between eye-modes and steady wind corkscrewing up through the eye has confirmed that positive feedback is generated, the condition for self-excited oscillation. Dimensions and speed of typical hurricanes require that the tones have frequencies on the order of $1/100$ of 1 Hz, which is doubtless the reason why they have escaped previous detection. Visual evidence appears in multiple alternating rain and dry bands that originate around the inner circumference of the eye wall, observable in aerial photos by NOAA. These are most easily explained as due to intense acoustically-driven temperature oscillations about the dew point, inside the eye. For a given hurricane, the actual tonal frequencies can be deduced from the spacing of these startups.

9:00

4aPAa4. Some medical applications of acoustic streaming. Lawrence A. Crum (Ctr. for Industrial and Medical Ultrasound, Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

Classical acoustic streaming results when an acoustic field is absorbed in a liquid, and the resulting momentum transfer causes the liquid to be translated. Acoustic microstreaming results when a bubble or other compressible entity is caused to oscillate by an acoustic field and hydrodynamic flow is induced in the vicinity of the oscillating surface. Much of the pioneering work on this general topic was performed by Wesley Nyborg and his students at the University of Vermont. There are a number of conditions for which acoustic streaming can be useful in medical ultrasound; e.g., one can enhance diffusion rates of drugs by inducing streaming at a specific site in tissue; together with Doppler imaging, one can determine the consistency of a particular sample of fluid, such as a blood clot. These and other examples of the use of streaming in medical ultrasound will be presented. [Work supported in part by the NIH and NSBRI.]

9:20

4aPAa5. The mechanical index and cavitation in tissue. Charles C. Church and Xinmai Yang (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677)

The mechanical index (MI) quantifies the likelihood that exposure to diagnostic ultrasound will produce an adverse biological effect by a nonthermal mechanism. The current formulation of the MI is based on inertial cavitation thresholds in two liquids, water and blood, as calculated by Apfel and Holland [Ultrasound Med. Biol. **17**, 179–185 (1991)]. Although tissue contains a high proportion of water, it is not a liquid but a viscoelastic solid. The importance of this difference was studied by deriving a Keller-Miksis-like equation assuming a gas bubble in a linear Voigt solid and performing numerical computations similar to the analytical work underlying the MI. Thresholds for inertial cavitation were determined for pulse lengths of 1–14 acoustic periods, equilibrium bubble radii (R_o) of 0.1–10.0 mm, a frequency range of 0.5–15.0 MHz, 2 threshold (P_t) criteria: $R_{max} = 2R_o$, and $T_{max} = 5000$ K, and tissue elasticities and viscosities of $G = 0, 0.5, 1.0$ or 1.5 MPa and $\mu = 0.005$ or 0.015 Pa·s; which span the range of values for soft tissue. It is found that thresholds in tissue are up to 10 times those in liquid, and that P_t increases nearly linearly with frequency. The relevance of these results to ultrasound safety will be discussed.

9:40

4aPAa6. Wesley Nyborg and bioacoustics at the University of Vermont. Junru Wu (Dept of Phys., Univ. of Vermont, Burlington, VT 05405)

Wes Nyborg came to UVM in 1960. Burlington, Vermont has become his home. He did his most pioneering research in microstreaming, acoustic radiation pressure and bioeffects of ultrasound at UVM. His research was continuously supported over 20 years by NIH. In 1987, after receiving Ph.D. in Physics and working with Isadore Rudnick for two years as a postdoc in the field of nonlinear physics, I joined the faculty of the University of Vermont as an assistant professor. Dr. Nyborg just retired from teaching in 1986. He was the person who convinced me to start research in the field of biomedical ultrasound. This presentation will review the development and achievement of bioacoustics at UVM and will also introduce him as a mentor, friend, colleague and collaborator.

Contributed Papers

10:00–10:30 Break

10:30

4aPAa7. A technique for monitoring and controlling cavitation activity during high intensity focused ultrasound application. Charles R. Thomas, Caleb H. Farny, Ronald A. Roy, and R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington Street, Boston, MA 02215)

It has been reported that cavitation can lead to enhanced heating in the focal region of a HIFU source. In order to exploit this heating for *in vivo* use, it is essential that the cavitation only occur in the focal region. Thus, the onset and evolution of inertial cavitation activity must be monitored and controlled during HIFU therapy. One candidate sensor is a confocally-aligned passive cavitation detector; however this would add complexity to a clinical HIFU applicator. Instead we propose that the HIFU source itself can serve as a monitoring device. The combination of broadband acoustic emissions (inertial cavitation) and backscatter (stable cavities) emanating from the HIFU focus manifests itself as fluctuations of the otherwise constant driving voltage amplitude, providing a convenient means for sensing cavitation activity. We will present results of experiments assessing the feasibility of using the variance in the amplitude of the HIFU drive voltage as a feedback control signal. Success is determined by the shape of the lesion created using supplemental control compared to the shape obtained otherwise. Two independent control parameters were used: the amplitude and the duty cycle of the HIFU. [Work supported by the Department of the Army award DAMD17-02-2-0014.]

10:45

4aPAa8. Spherical bubble pulsation between parallel plates. Jianying Cui, Mark F. Hamilton (Dept. of Mech. Eng., Univ. of Texas, 1 University Station C2200, Austin, TX 78712-0292), Preston S. Wilson, and Evgenia A. Zabolotskaya (Univ. of Texas, Austin, TX 78713-8029)

A simple model is proposed for investigating bubble pulsation in a constrained medium. In free space, compressibility of the surrounding medium introduces primarily radiation losses that can often be neglected. In a constrained medium, compressibility plays an essential role. The example we investigate here is linear pulsation of an acoustically driven spherical bubble between two parallel rigid plates. The method of images is used to obtain a bubble dynamic equation that accounts for time delays associated with reflections of pressure waves radiated by the bubble. When the bubble is midway between the plates, the series of terms for the reflections can be expressed in closed form. For narrow plate separations, less than about ten bubble diameters, the bubble resonance frequency is reduced from its value in a free field due to increased effective inertia of the liquid, and radiation damping is increased. For larger plate separations, but less than the acoustic wavelength, the quality factor associated with radiation damping is proportional to plate separation. If effects of compressibility are ignored, the effective inertia becomes infinite, and bubble pulsation is prohibited by the equations of fluid mechanics. [Work supported by NIH Grant EB004047 and ARL IR&D funds.]

11:00

4aPAa9. Evaluation of backscattered intensity to quantify the destruction rate of echogenic liposomes. Tyrone M. Porter, Sampada S. Vaidya, Christy K. Holland (Dept. of Biomed. Eng., Univ. of Cincinnati, Cincinnati, OH 45267), Shao-Ling Huang, Robert C. MacDonald (Northwestern Univ., Evanston, IL), and David D. McPherson (Northwestern Univ., Chicago, IL)

Echogenic liposomes (ELIP) are vesicles with a phospholipid bilayer shell that can serve as ultrasound contrast agents (UCA) for diagnostic purposes and for targeted drug delivery. The efficacy of ELIP as an UCA depends upon its stability in an acoustic field, whereas the use of ELIP for drug delivery will require releasing the encapsulated drug rapidly at the desired treatment site. The objective of this study was to evaluate the rate of destruction of ELIP as a function of pressure and pulse repetition frequency (PRF). Assuming ELIP destruction is directly related to its echogenicity, transducers with center frequencies of 3.5, 7.5, and 10 MHz were used to acquire pulse-echo data from ELIP suspensions ([lipid] = 0.2 mg/ml). The rate of ELIP destruction was defined by fitting an exponential decay function Ae^{-kt} to the backscattered intensity data. The relationship between ELIP destruction and acoustic pressure and PRF was determined by comparing the decay time constant k across samples. Based on this analysis, the rate of ELIP destruction was found to be directly proportional to pressure and PRF for all frequencies tested.

11:15

4aPAa10. Microbubble oscillations and stability for drug delivery. John S. AllenIII (Dept. of Mech. Eng., Univ. of Hawaii-Manoa, 2540 Dole St., Honolulu, HI 96822)

Ultrasound contrast agents have been developed from micron size bubbles whose gas core is enclosed by a polymer, lipid or protein shell. Furthermore, specific designs have been developed for drug delivery in which the ultrasound contrast agent acts as drug delivery vehicle. A drug may be suspended in the shell of these agents which is released at a

particular site as the microbubble undergoes destruction. Localized delivery depends on the destruction of a sufficient number of bubbles within a confined geometry typically a capillary or small vessel. Experimental evidence suggests the composition and thickness of the shell play important role in the break-up. Break-up scenarios of buckling have been suggested for elastic shell agents and shape stability for fluid shell agents. However, these have not been rigorously examined theoretically. The formulations of contrast agent designed for drug delivery are highlighted including a description of a double polymeric layer design. Shape stability equations for fluid shell agents are derived and analyzed in limiting cases. Furthermore, biological issues related to potential endothelial cell transport and interactions are briefly discussed.

11:30

4aPAa11. Cavitation inception on micro-particles: Possible drug carriers? Manish Arora, Bram Borkent, and Claus-Dieter Ohl (Phys. of Fluids, U. Twente, Postbus 217, 7500 AE Enschede, The Netherlands)

Hydrophobic particles are known to act as cavitation nuclei when they are exposed to a sufficient tensile wave. Yet, the dynamics of cavitation inception from particles, the growth of the bubble and the separation from the particle has only recently been observed. [M. Arora, C. D. Ohl, and K. A. Moerch, Phys. Rev. Lett. **92**, 174501 (2004)] In this presentation, high-speed photography of the remarkable dynamics is presented together with a modeling effort using a force balance approach. One of the results is that particles are accelerated to velocities up to 10 m/s and detach from the bubble. It was proposed, that micro-particles might act as drug carriers which can be activated by a tensile wave and accelerate into tissue or cells. In an effort to explore this possibility experiments on the cavitation inception ability on various types of micro-particles have been conducted in a reproducible way. Cavitation inception is initiated with a single cycle tensile wave and recorded with a camera. The number of cavitation bubbles decreases from shot-to-shot, which can be explained with one-time trigger-able nucleation sites on the particles.

11:45–12:00
Open Comments

THURSDAY MORNING, 19 MAY 2005

PLAZA A, 8:00 TO 10:15 A.M.

Session 4aPAb

Physical Acoustics: Wind Noise and Atmospheric Sound Propagation

Richard Raspet, Chair
Univ. of Mississipi, NCPA, Coliseum Dr., University, MS 38677

Contributed Papers

8:00

4aPAb1. Framework for windnoise studies. Richard Raspet, Jeremy Webster, and Kevin Dillion (Dept. of Phys. and Astron., Univ. of Mississippi, P.O. Box 1848, University, MS 38677)

Research in wind noise reduction in outdoor measurement microphones has been limited largely to comparisons between bare and screened microphones. Morgan and Raspet [J. Acoust. Soc. Am. **92**, 1180–1183 (1992)] used simultaneous wind velocity and noise measurements to show that the source of wind noise is incident wind fluctuations. In this paper, two methods for predicting the upper limits of wind noise pressure spectra

from velocity spectra in the inertial range are developed. A lower limit on wind noise is estimated from two theories of the intrinsic turbulent pressure fluctuations. Empirical results for the self-noise windscreens in substantially non-turbulent flows are also presented. Measurements of the wind velocity spectra and wind noise spectra from a variety of windscreens are described and compared to the theoretical predictions. All of the wind noise data lies between the upper and lower limits. The theoretical framework allows windscreens to be evaluated in terms of the best and worst-case scenarios and establishes practical lower limits on wind noise reduction for varying wind conditions. [Work supported by the Collaborative Technology Alliance sponsored by the US Army Research Laboratory.]

4a THU. AM

8:15

4aPAb2. Nighttime traffic noise in an urban environment. Kenneth E. Gilbert, Roger Waxler, Carrick L. Talmadge, and James P. Chambers (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677, kgilbert@olemiss.edu)

Traffic noise ducted near the ground at night and in the early morning can be represented simply in terms of normal modes. The effects of surface loss and atmospheric absorption, as well as meteorology, are contained in the mode attenuation coefficients. For a long (5–10 km) road section, the average noise levels are a function of the product of the mode attenuation coefficients and the perpendicular distance from the roadway. This paper considers a two-dimensional model for computing the effective mode attenuation coefficients in an urban environment. Modal attenuation is estimated by propagating modes over an irregular surface that approximates typical urban structures. The mode attenuation due to scattering from the urban structures is compared to the attenuation from atmospheric absorption and from a finite ground impedance. Predicted noise levels are compared to available data and the implications for nighttime urban traffic noise are discussed.

8:30

4aPAb3. The stability of the nocturnal quiet height. Roger Waxler, Kenneth E. Gilbert, (NCPA, U. of Mississippi, University, MS 38677, rwax@olemiss.edu), and Carrick Talmadge (NCPA, University, MS 38677)

It has recently been demonstrated that for narrow band signals propagating in the nocturnal boundary layer there is a height, a few meters off the ground, at which the sound pressure is significantly reduced. In this presentation the stability of the quiet height against fluctuations of the sound speed will be discussed. It is demonstrated theoretically, and verified by experiment, that, in the first few meters of the atmosphere, the dependence of the sound pressure level on altitude, relative to the sound pressure level at any fixed altitude, is insensitive to sound speed fluctuations.

8:45

4aPAb4. Wind noise at a flush microphone in a flat plate. Kevin Dillion, Richard Raspet, and Jeremy Webster (Dept. of Phys. and Astro., Univ. of Mississippi, P.O. Box 1848, University, MS 38677)

Elliot [Boundary Layer Meteorology 2, 476–495 (1972)] states that a flush-mounted microphone can measure the turbulent pressure fluctuations in outdoor flows without self-noise problems. This suggests that flush mounted microphones can be used to minimize wind noise. Wind velocity spectra from a hot wire anemometer and wind noise spectra from a flush-mounted microphone in a flat plate at ground level have been measured outdoors. The velocity power spectral densities and average velocities as a function of height are determined from the data. The measured pressure power spectral densities are compared to theoretical values calculated from fluid dynamic and meteorological turbulence theory. In particular we investigate the contributions of predicted turbulence-turbulence and turbulence-mean shear contributions to the pressure spectrum measured at the surface of the plate. [Prepared in part through collaborative participation in the Collaborative Technology Alliance for Advanced Sensors sponsored by the US Army research Laboratory under Cooperative Agreement DAAD19-01-0008.]

9:00

4aPAb5. Spatial correlations in fluctuations induced by long-range propagation of sound in the atmosphere. Carrick L. Talmadge, Shantharam Dravida, Kenneth E. Gilbert, and Roger Waxler (Univ. of Mississippi, NCPA, Oxford, MS 38677, clt@olemiss.edu)

Amplitude and phase fluctuations associated with sound propagating over long distances are well known to increase rapidly with distance [e.g., Wilson, Noble, and Coleman, J. Atmos. Sci. 60, 2473 (2003)]. This increase in the fluctuation level is usually given as an argument that the

sound source becomes temporally incoherent after propagating large distances, making signal processing techniques that rely on the temporal coherence of the signal impracticable for single microphone measurements. Results are presented from a series of experiments designed to measure the spatial coherence of sound in both day- and nighttime conditions. In these experiments, 24-element horizontal and 8-element vertical arrays were used to measure the coherence of sound sources over the frequency range 30–200 Hz (horizontal array) and 100–500 Hz (vertical array). In many of these experiments, surprisingly large spatial coherences in the source strength were observed. These results suggest that array measurements may allow us to compensate for the poor temporal coherence of the signals for long-range propagated sound. We will also discuss the different sources of the observed fluctuations in the sound level and phases for daytime versus nighttime conditions. These sources include atmospheric turbulence, changes in the nocturnal boundary layer height and nocturnal gravity waves.

9:15

4aPAb6. Numerical simulation of acoustic tomography of the atmosphere. Vladimir E. Ostashev, Sergey N. Vecherin, George H. Goedecke (Phys. Dept., New Mexico State Univ., Las Cruces, NM 88003, vostashe@nmsu.edu), D. Keith Wilson (U.S. Army Engineer Res. and Development Ctr., Hanover, NH 03755), Alexander G. Voronovich (NOAA/ETL, Boulder, CO 80305), and Edward G. Patton (Natl. Ctr. for Atmospheric Res., Boulder, CO 80307)

A state-of-the-art array for acoustic travel time tomography of the atmosphere is under construction by several organizations in the U.S. The array will allow the estimation of the temperature and wind velocity fields within a tomographic volume located a few meters above the ground with a horizontal size of about 80 m. This paper is devoted to numerical simulation of acoustic travel time tomography of the atmosphere. The temperature and velocity fields within the tomographic volume are modeled with the use of quasi-wavelets and Large Eddy Simulation. Then, the travel times of sound propagation between different pairs of sources and receivers of the tomography array are calculated. Given these travel times, three algorithms for reconstruction of the temperature and velocity fields are employed: the cell approach, basis function approach, and stochastic approach. It is shown that the first of these approaches allows us to reliably estimate the mean temperature and velocity vector within the tomographic volume, while the other two give a good reconstruction of fluctuations in the temperature and velocity with respect to their mean values. [Work supported by ARO, Grants DAAD19-03-1-0104 and DAAD19-03-1-0341.]

9:30

4aPAb7. Modeling pulse propagation in the nocturnal boundary layer. Roger Waxler, Kenneth E. Gilbert (NCPA, Univ. of Mississippi, University, MS 38677, rwax@olemiss.edu), Sergey Kulichkov (Russian Academy of Sciences), and Carrick Talmadge (NCPA, University, MS 38677)

The propagation of broad band (0 to 1000 Hz) pulses in downward refracting atmospheres over lossy ground is modeled in the time domain by Fourier synthesis using a modal model for sound propagation in the frequency domain. Of particular interest is the dispersion of the pulse as it propagates. A conjecture of Chunchuzov, Bush, and Kulichkov [J. Acoust. Soc. Am. 88, 455–461 (1990)] that at long ranges such pulses have a universal form, is shown to be true. In general, at long ranges from the source, the pulse develops a narrow band tail, centered around 20 or 30 Hz. This tail is formed by the superposition of surface modes.

9:45

4aPA8. Prediction of aerodynamically generated sound via 3D particle velocity measurements. Sean Wu, Aditya Kumar, Zhi Ni (Dept. of Mech. Eng., Wayne State Univ., 5050 Anthony Wayne Dr., Detroit, MI 48202), and Hans-Elias de Bree (Microflow Technologies, AH Zevenaar, The Netherlands)

It has been shown [Wu and Hu, *J. Acoust. Soc. Am.* **103**, 1763–1774 (1998)] that acoustic radiation from a vibrating object can be calculated directly by an Alternate Integral-formulation Method (AIM), once the particle velocity distribution over a hypothetical surface enclosing this object is specified. This concept is extended here to prediction of aerodynamically generated sound. The particle velocity distribution on a hypothetical enclosure is measured by a fully integrated sound pressure-velocity ($p-u$) probe. This novel $p-u$ sensor combines a MEMS particle velocity sensor and miniature pressure transducer, which allows for measurements of the acoustic pressure and particle velocity at a single point simultaneously. It is simple and easy to use and is suitable for both narrow and broadband sound measurements. This $p-u$ probe is employed to measure the particle velocity field generated by a typical hairdryer. Measurements thus obtained are taken as input to AIM to predict the radiated acoustic pressure field. Experiments are conducted inside an anechoic chamber at the Acoustics, Vibration, and Noise Control Laboratory at Wayne State Uni-

versity. The predicted acoustic pressures are then compared with those measured at the same locations by microphones. Satisfactory agreements are obtained at all frequencies.

10:00

4aPA9. A comparison of conventional and metal foam windscreens. Gunnar R. Becker, Guenther H. Hermstruewer, and Rainer Knoetsch (Rheinmetall Defence Electron. GmbH, Brueggeweg 54, 28309 Bremen, Germany)

In outdoor applications wind induced noise reduces the signal quality significantly. Conventional windscreens made of polyurethane (PU) foam are widely used to reduce this noise. However under harsh environmental conditions such as high temperatures, exposure to UV radiation, etc., PU may suffer. An alternative may be the usage of metal foams instead of PU foams. Spherical windscreens made of open-pored metal foam with dimensions close to commercial ones and different pore distributions have been manufactured by m-pore and were tested under different wind conditions. Comparative measurements between both conventional PU- and metal foam windscreens indicated that the reduction of wind induced noise is almost the same. A reduction of the receiver level due to the windscreens was not observed. Measurements and test results will be presented. [Work supported by m-pore.]

THURSDAY MORNING, 19 MAY 2005

PLAZA A, 10:45 A.M. TO 12:00 NOON

Session 4aPac

Physical Acoustics: Thermoacoustics and Particle Agglomeration

Matthew E. Poese, Chair

Applied Research Lab., Pennsylvania State Univ., State College, PA 16804

Contributed Papers

10:45

4aPac1. Specific acoustic impedance measurements of a thermoacoustic stack. Heui-Seol Roh, Richard Raspet, and Henry E. Bass (Dept. of Phys. and Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677)

Simmons [Ph.D. dissertation, University of Mississippi, 2003] attempted to use a two-microphone specific acoustic impedance (SAI) measurement of a thermoacoustic stack to determine the thermoviscous functions and the coefficient of the thermal gradient in the acoustic pressure differential equation for arbitrary geometry stacks. The proposed method was tested in part by comparing the measured and predicted SAI of a well-characterized parallel pore stack. Significant disagreement between theory and measurement occurred. In this paper the results of an improved four microphone SAI measurement on the well-characterized stack are reported. This measurement will reduce uncertainties in the transfer matrix method used by Simmons and determine if the discrepancies in Simmons work were experimental or if the thermoacoustic theory is incomplete.

11:00

4aPac2. High frequency operation of thermoacoustic coolers and prime movers. Husam El-Gendy, Young Sang Kwon, and Orest G. Symko (Dept. of Phys., Univ. of Utah, 115 S. 1400 E., Rm. 201, Salt Lake City, UT 84112-0830)

By operating thermoacoustic engines at high frequencies, 4 kHz and higher, the devices have characteristics which are important for many applications. Since they are resonant systems, the power density increases with frequency. Reduction of device size provides quick thermal response time in both the cooler and the prime mover. Moreover, small device size

makes it practical to incorporate them into arrays, which can handle large powers. Most important is the fact that small devices make it simple for operation at high pressures in working gas without exceeding strength of materials limitations. This leads to high power densities. Results will be presented to illustrate how the above features affect device performance for the frequency range of 4 kHz to 21 kHz. Measurements using Particle Image Velocimetry of streaming, instabilities, and resonator mode interactions will be discussed for this high frequency range. Ultimately as the operating frequency is raised, device efficiency is limited by heat conduction along the stack and working fluid. [Work supported by the Office of Naval Research and the State of Utah.]

11:15

4aPac3. Initial investigations into thermoacoustic coal agglomerators. Gordon Smith (Dept. of Phys. and Astronomy, Western Kentucky Univ., 1 Big Red Way, Bowling Green, KY 42101)

Each year, Kentucky coalmines produce about 160 million tons of coal, mainly used to generate electricity at power plants in Kentucky and throughout the midwestern and southeastern United States. Coal processing generates an excessive amount of pollution in the form of heavy exhaust, which creates an environmental concern. In response to this growing concern for a balance between environmental impact and energy production, many processing plants utilize electrostatic precipitation (ESP) systems to agglomerate and filter out particulate matter (PM) from exhaust streams before they enter the general environment. However, ESP efficiency is limited by the size of the particulate exhaust. Particle agglomeration also results via time-averaged forces inherent in an acoustic standing wave. Thermoacoustic technology, which utilizes waste heat sources for operation, embodies an ideal source for providing the requisite acous-

4a THU. AM

tic conditions. This technology could also be installed to supplement existing technology as a preconditioner (increasing the particle size to one more suitable for current ESP systems), and increase the overall PM removal efficiency. This talk will present initial investigations into applying thermoacoustic technology to this problem.

11:30

4aPac4. Experimental study of aerosol concentration in flow-through, low-frequency resonators. Douglas Meegan, Justin Smith, and Wayne Wright (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029)

An experimental method to study the concentration of micron-sized aerosols in a low-frequency sound field was developed. The physical mechanism for acoustic aerosol concentration in a low-frequency sound field is the asymmetric Stokes drag which has been predicted to exceed the effects of radiation pressure for typical aerosol sizes [Meegan and Ilinski, *J. Acoust. Soc. Am.* **114**, 2387 (2003)]. Related experiments were conducted in which 5 micron aerosol was drawn through a small duct that was driven in a transverse resonant mode with peak sound pressure levels of greater than 150 dB re 20 microPascals and frequencies in the range 1 to 10 kHz. The aerosol stream was illuminated by a laser sheet through the transparent walls of the duct in order to visualize (by video) the effects of the sound field. The experiments confirm the basic feature of the asymmetric Stokes drag model—specifically, the aerosol is observed to concen-

trate along acoustic velocity nodes. Ongoing experiments to quantify the level of concentration will be described. [Work supported by RDECOM.]

11:45

4aPac5. The effects of orthokinetic collision and the acoustic wake effect on acoustic agglomeration of polydisperse aerosols. Shaozeng Dong (Dept. of Mech. Eng., Virginia Commonwealth Univ., 601 West Main St., Richmond, VA 23284), Bart Lipkens (Western New England College, Springfield, MA 01119), and Timothy Cameron (Kettering Univ., Flint, MI 48504)

A new concept of effective agglomeration length, which measures the maximum particle separation distance for effective collisions, is proposed for acoustic agglomeration of polydisperse aerosols with respect to the separate and combined effects of orthokinetic collision and acoustic wake in a horizontal acoustic wave. Particle gravity is found to be significant for the acoustic wake effect while the particle collision efficiency is important for the orthokinetic collision. Results indicate that orthokinetic collision dominates at low frequencies for intermediate size ratios while the acoustic wake effect is more significant at higher frequencies for all particles. The optimum frequency for orthokinetic collision is confirmed but shifts downward with the increase of sound power. For the acoustic wake effect, the agglomeration increases monotonically with sound frequency. Results also show that the orthokinetic collision is not effective for agglomeration of sub-micron particles because of low particle collision efficiency.

THURSDAY MORNING, 19 MAY 2005

REGENCY C, 8:00 A.M. TO 12:00 NOON

Session 4aPP

Psychological and Physiological Acoustics: Temporal Factors, Masking and Pitch (Poster Session)

Magdalena Wojtczak, Chair

Dept. of Psychology, Univ. of Minnesota, 75 East River Rd., Minneapolis, MN 55455

Contributed Papers

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon.

4aPP1. The temporal effect for signal frequencies around a notched cochlear hearing loss. Elizabeth Strickland and Lata Krishnan (SLHS Dept., Purdue Univ., W. Lafayette, IN 47907)

The temporal effect (or overshoot) refers to the change in signal-to-masker ratio at threshold for a short-duration tone as it is delayed from the onset of a longer-duration masker. This study is one in a continuing series of studies on the relationship between the temporal effect and the cochlear active process. In a previous study, the temporal effect was measured as a function of signal level in a broadband masker, for listeners with cochlear hearing loss. Results showed that the pattern of the temporal effect with level depended not only on the degree of hearing loss at the signal frequency, but also on the degree of hearing loss above the signal frequency. In the present study, this was examined further by measuring the temporal effect as a function of signal level in listeners with notched cochlear hearing loss, when the signal frequency was below, at or above the frequency of the greatest hearing loss (the notch). Results will be analyzed in terms of changes in the amplification due to the cochlear active process at and above the signal frequency. [Work supported by the Kinley Trust.]

4aPP2. Effect of frequency on the relationship between intensity discrimination and the detection of amplitude modulation. Rebecca E. Millman and Sid P. Bacon (Dept. of Speech and Hearing Sci. Arizona State Univ., P.O. Box 870102, Tempe, AZ 85287-0102, rebecca.millman@asu.edu)

Wojtczak and Viemeister [*J. Acoust. Soc. Am.* **106**, 1917–1924 (1999)] found a clear relationship between the detection of intensity increments and the detection of amplitude modulation (AM) for continuous 1-kHz sinusoids. In the present study AM detection thresholds and intensity difference limens (DLs) were measured for 1-s gated sinusoids with a frequency of 125 Hz or 1 kHz. The sinusoids were presented at levels between 10 dB SL and 90 dB SPL. Modulation detection thresholds [$20 \log(m)$] improved as the carrier level was increased; for a comparable range of levels (45 dB), the improvement was similar at the two frequencies. Intensity DLs [$10 \log(\Delta I/I)$] generally improved as the level of the reference sinusoid was increased, although the improvement (“near miss” to Weber’s Law) was smaller at 125 Hz than at 1 kHz over the same 45-dB range of levels. For the 1-kHz sinusoids, the relationship between intensity DLs and modulation detection thresholds was consistent with that shown by Wojtczak and Viemeister (1999). This same relationship apparently does not hold at 125 Hz. Because the relationship between intensity dis-

crimination and modulation detection depends upon frequency, the mechanisms underlying the two tasks may be somewhat different. [Work supported by NIDCD.]

4aPP3. Level effects in amplitude modulation tuning. Magdalena Wojtczak and Neal F. Viemeister (Univ. of Minnesota, 75 East River Rd., Minneapolis, MN 55455, wojtc001@umn.edu)

Psychophysical experiments on amplitude-modulation (AM) masking suggest the existence of channels selectively sensitive to different ranges of modulation rates. This study investigates whether tuning depends on carrier level. Masking of a 40-Hz signal AM was measured as a function of the masker-AM rate selected from the range between 4 and 256 Hz, for Gaussian-noise and 5.5-kHz carriers. For the noise carrier, similar patterns were observed at two different spectrum levels (25 and 40 dB SPL), leading to similar estimates of the modulation-filter bandwidths. For the tonal carrier, however, tuning in AM masking varied considerably across levels. At the lowest level tested (40 dB SPL), the patterns were very broad and highly asymmetric. The patterns became more similar across carrier levels when a highpass noise was presented with the carrier. The noise limited the use of spread of excitation and thus, raised thresholds for the unmasked signal detection. The data suggest that the AM selectivity is strongly affected by the detection threshold for the unmasked AM. [Work supported by NIH grants DC00683 and DC006804.]

4aPP4. Comparison of intensity discrimination, increment detection, and comodulation masking release in the envelope and audio-frequency domains. Paul C. Nelson (Dept. of Biomed. & Chemical Eng. and Inst. for Sensory Res., Syracuse Univ., 621 Skytop Rd., Syracuse, NY 13210, pcnelson@syr.edu), Stephan D. Ewert (Tech. Univ. of Denmark, 2800 Kgs. Lyngby, Denmark), Laurel H. Carney (Syracuse Univ., Syracuse, NY 13210), and Torsten Dau (Tech. Univ. of Denmark, 2800 Kgs. Lyngby, Denmark)

In the audio-frequency domain, the envelope apparently plays an important role in detection of intensity increments and in comodulation masking release (CMR). The current study addressed the question whether the second-order envelope (“venelope”) contributes similarly for comparable experiments in the envelope-frequency domain. One set of experiments examined the relationship between gated intensity discrimination and continuous-carrier increment detection. In contrast to the asymmetry observed in the audio-frequency domain (listeners are more sensitive to increments), AM-depth discrimination thresholds were found to be the same in conditions with a continuous (modulated) carrier and with traditional gated stimuli for AM frequencies ranging from 4–64 Hz. The second set of experiments compared the amount of CMR in a tone-in-noise detection task when slow, regular fluctuations were imposed on the masking waveform in both domains. A significant release from masking of a 32-Hz signal in the modulation frequency domain was obtained only when the venelope fluctuations were slower than 1–2 Hz. Both experiments suggest a relatively weak contribution of venelope cues in the AM domain when compared to those provided by envelope cues in the spectral domain. [Work supported by NIH-NIDCD R01001641 (PCN, LHC) and the Danish Research Council (SDE, TD).]

4aPP5. Temporal integration functions of amplitude modulation detection and amplitude modulation depth discrimination. Jungmee Lee and Glenis Long (Speech and Hearing Sci., The Grad. School and Univ. Ctr., The City Univ. of New York, 365 5th Ave., New York, NY 10016)

Previous studies (Lee and Green, 1994; Lee and Bacon, 1997) suggested that both AM rate and AM depth discrimination were influenced by the number of AM cycles, instead of the duration of stimuli. AM detection and AM depth discrimination (standard depth=0.1) were measured as a function of the number of AM cycles for modulation rates of 10, 20, 40,

80, 125, 160, and 320 Hz. Different numbers of modulation cycles were used for each modulation rate: 2, 4, or 8 for 10 Hz; 2, 4, 8, or 16 for 20 Hz; 2, 4, 8, 16, or 32 for 40 Hz; 2, 4, 8, 16, 32, or 64 for 80, 125, 160, and 320 Hz. The carrier was a broadband-noise (10 kHz lowpass), and the carrier was either gated with the modulator or presented 250 ms earlier and 250 ms later than the modulator. The overall level of each presentation was randomized within 6-dB range from 65 dB SPL. The results suggest that there might be different temporal integration processes for AM detection and AM depth discrimination. The pattern is different for lower and higher modulation rates. [Work was supported by NIDCD Grant No. R03 DC06605-01.]

4aPP6. Coherent modulation enhancement: Improving performance in noise for hearing aids and cochlear implants. Pamela Souza (Dept. of Speech and Hearing Sci., Univ. of Washington, 1417 NE 42nd St., Seattle, WA 98105), Les Atlas, Steven Schimmel (Univ. of Washington, Seattle, WA 98105), Jay Rubinstein, Ward Drennan, and Jong Ho Won (Univ. of Washington, Seattle, WA 98105)

Difficulty hearing in noise is a pervasive problem for hearing-impaired listeners. One approach is to digitally reduce noise using a modulation filter, which can selectively modify the time envelope of selected spectral bands. Previous approaches to modulation filtering, such as those based upon a Hilbert transform magnitude, increased signal distortion. A new coherent approach was used to determine a fixed and an optimal adaptive modulation filter. The optimal adaptive filter used information from the target speech to design the modulation filter. Threshold signal-to-noise ratio was measured adaptively using a spondee-in-noise task [Turner *et al.*, *J. Acoust. Soc. Am.* **115**, 1729 (2004)]. Subjects selected the spondee heard in a forced-choice task with a two-talker babble background. The spondee level was fixed at the listeners most comfortable level and noise level adjusted adaptively using a 2 dB step size, with threshold based on 10 reversals. Sensorineural hearing-impaired listeners completed three conditions: unprocessed speech; fixed modulation filter; and optimal adaptive modulation filter. Normal-hearing subjects heard a parallel set of conditions processed to simulate a 6-channel cochlear implant. There was significant improvement in threshold signal-to-noise ratio with the optimal adaptive filter for both groups. There was no improvement for the fixed modulation filter relative to unprocessed speech.

4aPP7. Factors affecting gap duration discrimination performance. John H. Grose, Joseph W. HallIII, and Emily Buss (Dept. Otolaryngol.-HNS, Univ. of North Carolina at Chapel Hill, 1115 Bioinformatics CB#7070, Chapel Hill, NC 27599-7070, jhg@med.unc.edu)

This investigation tests the hypothesis that reduced acuity in temporal tasks is evident relatively early on in the aging process for tasks involving brief stimuli or intervals, across-frequency processing, and/or significant processing loads. In an earlier study of gap duration discrimination (GDD) using fixed 20-ms markers, it was found that young listeners performed significantly better than middle-aged listeners for both within-channel and across-channel configurations. However, this age effect did not emerge in a subsequent study that employed random-duration markers. The purpose of the present study was to clarify stimulus conditions under which early aging effects are evident in GDD tasks. Normal-hearing young (20–25 years) and middle-aged (40–55 years) listeners were tested using both fixed- (20-ms) and random- (40 ms ± 50%) duration markers in configurations that were either within-channel or across-channel. In addition, task complexity was manipulated by incorporating a rhythm discrimination feature into the GDD task. Results to date indicate that performance declines when marker duration varies randomly, both for within- and across-channel conditions, as well as when task complexity is increased. These findings will be discussed in the context of identifying early aging effects in temporal processing. [Work supported by NIDCD DC001507.]

4aPP8. Fundamental frequency discrimination and harmonic resolvability in sensorineural hearing impairment. Joshua G. W. Bernstein and Andrew J. Oxenham (MIT Res. Lab. of Electron. and Harvard-MIT Speech & Hearing Bioscience & Technol. Prog., 77 Massachusetts Ave., Cambridge, MA 02139, jgbern@mit.edu)

Sensorineural hearing loss (SNHL) often results in impaired fundamental frequency (F_0) processing. Three experiments tested the hypothesis that this deficit is related to a loss of frequency selectivity, which may result in fewer peripherally resolved harmonics. F_0 difference limens (F_0 DLs) were measured as a function of F_0 for bandpass filtered sine- and random-phase harmonic complexes in listeners with mild-to-moderate SNHL. All listeners showed a transition between small (good) F_0 DLs at high F_0 s and large (poor) F_0 DLs at low F_0 s, although the transition point varied across subjects. Two measures that are thought to reflect frequency selectivity generally corresponded to the F_0 DL transition point: the maximum F_0 for which F_0 DLs were phase-dependent, and the minimum modulation frequency required to discriminate amplitude modulation from quasi-frequency modulation. These results provide some support for the idea that the F_0 -processing deficit associated with SNHL is related to poor frequency selectivity. However, contrary to our hypothesis, poor correspondence was observed between the F_0 DL transition point and auditory filter bandwidths estimated using the notched-noise method in the same subjects. This suggests a possible discrepancy between traditional spectral-masking estimates of frequency selectivity and estimates based on temporal interactions between unresolved components. [Work supported by NIH grants R01DC05216 and 5T32DC00038.]

4aPP9. Modeling comodulation masking release using an equalization cancellation mechanism. Tobias Piechowiak, Stephan D. Ewert, and Torsten Dau (Ctr. for Appl. Hearing Res., Oersted-DTU, Tech. Univ. of Denmark, Bldg. 352, 2800 Kgs. Lyngby, Denmark, tda@oersted.dtu.dk)

Across-filter comparisons of temporal envelopes are a general feature of auditory pattern analysis which plays an important role in extracting signals from noise backgrounds. One example where a comparison of the temporal envelope in different frequency regions can lead to a substantial facilitation of detecting a signal in noise is comodulation masking release (CMR). It has been suggested that in (monaural) across-channel processing, like in (binaural) across-ear processing, an equalization-cancellation (EC) based circuit might be an effective strategy of noise reduction. The first part of the study investigates the relation between CMR and envelope-based binaural masking level differences (BMLD), using narrowband noise maskers and classical across-channel configurations (like N0Spi, N0Sm). In the second part, a model is presented that explicitly simulates CMR whereby the EC mechanism is assumed to take place at the output of a modulation filterbank. In the case of several flanker bands, the equalization is realized effectively by averaging the flanker bands, and cancellation by subtracting the averaged representation from the signal band. A generalized version of the model is presented that considers both across-channel and within-channel contributions to signal detection in CMR. [Work supported by the Danish Research Council.]

4aPP10. A unified view of the temporal-window and the adaptation-loop model in conditions of forward- and simultaneous masking. Ole Hau, Stephan D. Ewert, and Torsten Dau (Ctr. for Appl. Hearing Res., Oersted DTU, Tech. Univ. of Denmark, Bldg. 352, 2400 Kgs. Lyngby, Denmark, se@oersted.dtu.dk)

Two different mechanisms have been discussed in the literature as a possible explanation for forward masking: persistence (or temporal integration) and adaptation. In this study, two well established models of temporal processing in the auditory system are compared in a unified modeling framework. The temporal-window model representing a temporal-integration mechanism and the adaptation-loop model as the representative for the adaptation mechanism. The unified modeling framework shared a compressive, non-linear auditory filter stage and a template-based (optimal detector) decision stage. Predictions from both models

were obtained in conditions of forward masking at 1 and 4 kHz for masker-to-signal gaps of 0 to 160 ms, and in conditions of simultaneous masking, where the signal was temporally overlapping with the masker. The signal was a 10- or 12-ms raised-cosine ramped pure tone and the masker a 200-ms broadband noise. It was found that when both models are adjusted to behave similarly, the two mechanisms can be considered as being essentially equivalent. In fact, the combination of integration and the signal-to-noise-ratio based detection criterion in the temporal-window model can be interpreted as adaptation. Suggestions for a unified model of auditory processing are presented. [Work supported by the Danish Research Council.]

4aPP11. An examination of temporal signal uncertainty for sequences of noise or random-frequency maskers. Lori Leibold, Donna Neff, and Walt Jesteadt (Boys Town Natl. Res. Hospital, 555 North 30th St., Omaha, NE 68131, leiboldl@boystown.org)

The effect of temporal uncertainty was examined on detection of a 1000-Hz tone presented simultaneously with one of a sequence of five masker bursts, using a 2-AFC, adaptive procedure. The signal and each masker burst were 100-ms (10-ms rise/fall), with no temporal overlap between masker bursts. Each masker burst was 60 dB SPL. Across conditions, maskers were broadband-noise (300–3000 Hz) or random-frequency multi-tonal maskers with 1, 2, or 10 components. Components for the random-frequency maskers were drawn from 300 to 3000 Hz, excluding a 160-Hz band around the signal. Thresholds for conditions with no temporal uncertainty (signal presented with the first, third, or fifth masker burst on all trials) were compared to performance with maximal temporal uncertainty (signal position varied on every trial). There was no effect of temporal uncertainty for the noise masker. For random-frequency maskers, performance was uniformly poor for both fixed- and random-position signals, except for some release from masking for signals presented with the last masker burst. For three of four listeners, even 1-component maskers showed large amounts of masking in all conditions. These results suggest that the effects of temporal uncertainty are minimal for these stimuli. [Work supported by NIDCD.]

4aPP12. Spectral integration and multiple looks. Robert Lutfi (Dept. of Communicative Disord. and Waisman Ctr., Univ. of Wisconsin, Madison, WI 53706)

A fundamental property of hearing is that signals become more detectable as their bandwidth is increased. The rate at which detection improves is reasonably well predicted by models that assume integration of energy at the output of a single auditory filter roughly matched in bandwidth to the signal. This paper tests an alternative account in which detection is mediated by not one but several independent auditory filters whose outputs are combined (multiple-looks model). Listeners detected an increment in the level of a multitone complex whose bandwidth was increased by adding tones at successively higher and lower frequencies. In different conditions the frequencies were fixed (F) or were perturbed at random from trial to trial with the perturbation being the same (PS) or different (PD) for each frequency. The multiple-looks model predicts performance should improve with bandwidth at a faster rate for PD and at a slower rate for PS compared to F; this due to differences in the shared variance in the output of separate auditory filters. The single filter model predicts little effect of frequency perturbation in either case. Individual differences in the results provided partial support for both models. [Work supported by NIDCD.]

4aPP13. Response growth with and without a low-frequency suppressor. Jeffrey J. DiGiovanni (School of Hearing, Speech and Lang. Sci., Ohio Univ., Athens, OH 45701) and Magdalena Wojtczak (Univ. of Minnesota, Minneapolis, MN 55455)

In a psychophysical study, Wojtczak and Viemeister (2005) demonstrated that the response to a tone suppressed by a fixed-level, higher-frequency suppressor grows faster than the response to the same but un-

suppressed tone. This finding is consistent with a linearization of the response under suppression. This study extends their experiment to suppressors with frequencies below that of the suppressee. Detection of a 10-ms, 4-kHz probe was measured under two forward-masking conditions: one with a 200-ms, 4-kHz masker presented alone, and the other with the same masker/probe paired with a 2.4-kHz fixed-level suppressor. A range of probe levels was used to measure growth of masking with and without the suppressor. The 4-kHz masker level was varied adaptively to find the masked threshold. Suppression was revealed by the difference between the masker levels needed for the masked threshold in the presence and absence of the suppressor. Initial comparison of the rates of masking growth between the two conditions suggests a linearization of the response to the 4-kHz masker in the presence of a lower-frequency suppressor. A few levels of the suppressor were used to determine the rate of suppression growth and its effect on the growth of response at the suppressee-frequency place.

4aPP14. Stochastic resonance in Gaussian and uniform noise. Dennis Ries (School of Hearing, Speech, and Lang. Sci., Ohio Univ., Grover Ctr. W221, Athens, OH 45701, ries@ohio.edu)

The improvement in threshold of a 2.0 kHz tone in the presence of low levels of uniform and Gaussian noise is compared to tonal threshold measured in quiet. Uniform noise was created using a rectangular amplitude distribution. The noises were low-pass filtered at 12.0 kHz and were presented continuously throughout a run at spectrum levels ranging from 0 to -30 dB/Hz. Pure-tone signals were 400 ms in duration including 10 ms raised cosine onset and offset ramps. Thresholds were obtained using a 3-interval, forced-choice procedure with correct answer feedback in conjunction with a two down, one up adaptive tracking paradigm that targets 70.7% correct on the psychometric function. Subjects were instructed to select the interval that was different from the others. Preliminary results for the Gaussian noise conditions are similar to earlier results [Zeng *et al.*, Brain Res. **869**, 251–255 (2000)]. Initial results for the uniform noise indicate that it produced equivalent or slightly lower thresholds than those measured in the presence of Gaussian noise.

4aPP15. Spatial release from informational masking along the front-back dimension. Neil L. Aaronson (Dept. of Phys. and Astron., Michigan State Univ., 4230 BPS Bldg., East Lansing, MI 48824), Brad Rakerd (Michigan State Univ., East Lansing, MI 48824), and William M. Hartmann (Michigan State Univ., East Lansing, MI 48824)

When two independent speech samples are presented together from a single location in front of the listener, one will mask the other. The amount of masking can be reduced by presenting a repeated masker from a different location off to the side and shifting it slightly forward (+) or backward (-) in time compared to the masker in front. New experiments, using the coordinate response measure technique with a two-female-talker masker and a female target, show that masking release can also be obtained when the target and masker are in front and the repeated masker is directly in back. Release is seen for both forward and backward time shifts, ranging from -32 to 32 ms. The amount of release is somewhat more than half that obtained when the repeated masker is off to the side. Release from masking can also be seen when the repeated masker comes from a location directly above the target and masker in front, but only for a single value of time shift, namely ± 2 ms. It is concluded that both

spatial and spectral cues mediate release from informational masking in the front-back dimension. [Work supported by the NIDCD grant DC 00181.]

4aPP16. The advantage of knowing where to listen. Gerald Kidd, Jr., Tanya L. Arbogast, Christine R. Mason, and Frederick J. Gallun (Commun. Disord. and Hearing Res. Ctr., Boston Univ., 635 Commonwealth Ave., Boston, MA 02215)

This study examined the importance of *a priori* knowledge about the location of a target in a multitalker environment. On each trial, three equal-level sentences from the Coordinate Response Measure test were presented from different loudspeakers separated by 60° azimuth. The sentences—target and 2 maskers—had the structure “Ready [callsign] go to [color] [number] now.” The task was to repeat the color and number associated with a specific callsign. The target location was chosen among the 3 loudspeakers on every trial. The main experimental variables were the probability of occurrence of the target at each speaker location and whether the listener was provided with the callsign of the target before or after stimulus presentation. Performance ranged from near perfect when the target location was completely certain to about 33% correct when the location was random and the callsign was not known in advance. When target location was random, performance improved from about 33% to 70% correct when the callsign was known in advance. Overall, these results support the view that knowing where to focus attention provides a great advantage in speech recognition in complex and uncertain environments. [Supported by NIH/NIDCD.]

4aPP17. Effect of signal frequency uncertainty for random multi-burst maskers. Rong Huang and Virginia M. Richards (Dept. of Psych., Univ. of Pennsylvania, 3401 Walnut St., Ste. 302C, Philadelphia, PA 19104)

The detectability of a sequence of equal-frequency tone pips masked by random multi-burst complexes may depend on the perceptual segregation of the signal stream from the random masker. If so, detection thresholds may be independent of whether the signal frequency is known versus uncertain. In this experiment observers detected a signal stream of 8 sequential equal-frequency 30 ms tone pips embedded in a random masker composed of 8 sequential bursts. A yes/no procedure was used, and the independent variable p (probability a tone was played at a particular time-by-frequency location) governed the number of masker tones in each burst. The dependent variable was d . Threshold values of p were obtained for signal streams at 5 different frequencies. Sensitivity was superior for the mid-frequency signal, and decreased as the signal frequencies approached the edge of the frequencies the masker tones occupied (200–5000 Hz). When the frequency of the signal stream was randomly varied from trial to trial, sensitivity was poorer than for any of the fixed-frequency signals. Thus, the detectability of a sequence of tone pips is reduced when the signal frequency is uncertain compared to certain. Additionally, sensitivity increased when the signal stream was delayed relative to the masker bursts.

4aPP18. Training listeners or preserving phase information improves the effect of perceived spatial separation on releasing spectrally degraded Chinese speech from information masking. Jing Chen, Chenfei Ma, Chun Wang, Hongwei Qu (Natl. Key Lab. on Machine Percept. Speech and Hearing Res. Ctr., Dept. of Psych., Peking Univ., Beijing 100871, China), Xihong Wu, Liang Li (Peking Univ., Beijing 100871, China), and Bruce Schneider (Univ. of Toronto at Mississauga, Canada L5L 1C6)

Physical or perceived spatial signal/masker separation unmasks speech more when maskers are informational than when energetic. However, it is unclear how beneficial the separations are to cochlear-implant listeners, because signal transductions applied in cochlear implant degrade signals

spectrally, and spectrally degraded speech is more vulnerable to maskers. Here, spectrums of both target speech (nonsense sentence) and masker (steady speech-spectrum noise, speech modulated speech C -spectrum noise, or speech) were filtered into 15 frequency bands. For both target and masking speech, the center-frequency pure tone of each band was modulated by the extracted envelope from the band. The target speech was composed by the sum of the 8 odd-band tones, and the masker was either same-band (with the 8 odd-band tones) or different-band (with the 7 even-band tones). The results show that physical but not perceived spatial separation unmasked target speech in naive normal-hearing listeners. However, following pre-presentations of both degraded and normal correspondent speech to listeners for a period of time or the introduction of phase information into modulated tones, perceived spatial separation reduced the influence of different-band speech masking but not that of same-band speech masking. These results are useful for improving cochlear-implant programs at both behavioral and technical levels.

4aPP19. Development of backward masking in elementary school children. Cynthia M. Zettler, Marsha G. Clarkson, Rose A. Sevcik, and Robin D. Morris (Dept. of Psych., Georgia State Univ., Atlanta, GA 30303)

Previous research on backward masking (BM) suggests that the ability develops slowly across childhood. When a spectral notch is placed in a noise masker, thresholds improve due to a decrease in masker energy at the frequency of the signal. To further clarify the developmental course of BM thresholds for BM and backward notched-noise (BM-N), conditions were measured in 81 children ranging in age from seven to ten years. BM stimuli consisted of a 20-ms, 1000-Hz tone presented 20 ms prior to a 300-ms, 600–1400 Hz bandpass noise. BM-N stimuli consisted of a 1000-Hz tone presented 20 ms prior to the onset of a 300-ms, 400–1600 Hz bandpass noise with a spectral notch between 800 and 1200 Hz. Multivariate analysis of variance indicated a significant difference among age groups for the BM condition ($F(3,77)=3.63, p<.05$), but not for the BM-N condition. Thresholds generally decreased with increasing age in the BM condition (age seven 87.86 dB; eight 81.36 dB; nine 75.31 dB, and ten 77.26 dB). Increased BM thresholds in younger children may indicate poorer processing efficiency, whereas the notched results suggest that frequency selectivity is relatively mature by age seven.

4aPP20. Tone-in-noise detection using narrowband reproducible maskers with restricted energy and envelope cues. Sean A. Davidson (Dept. of Biomed. and Chem. Eng., Inst. for Sensory Res., 621 Skytop Rd., Syracuse Univ., Syracuse, NY 13244, sadavids@syr.edu), Robert H. Gilkey (Wright State Univ., Dayton, OH), and Laurel H. Carney (Syracuse Univ., Syracuse, NY 13244)

Both energy- and envelope-based models have been successfully used to predict narrowband tone-in-noise detection thresholds. To distinguish between these models, hit and false-alarm rates for 25 reproducible maskers were measured in the N_0S_0 and N_0S_π interaural configurations and 4 cue conditions: (1) normal energy and envelope cues, (2) restricted envelope cues, (3) restricted energy cues, or (4) restricted energy and envelope cues. Preliminary results show that hit and false-alarm rates were strongly correlated between the conditions with and without restricted energy cues, indicating that the reduction of energy cues *did not* substantially affect the detection process. However, hit and false-alarm rates were more weakly correlated between the conditions with and without restricted envelope cues, suggesting that the reduction of envelope cues *did* affect the detection process. Restricting envelope cues also had a larger (but still modest) effect on d' under the N_0S_π configuration, but restricting energy cues had a larger effect for the N_0S_0 configuration. Overall, these results are more consistent with envelope-based models, but indicate that additional cues must play a role. [Work supported by NIDCD R01-DC-001641 (LHC, SAD) and the Ohio Board of Regents (RHG).]

4aPP21. The spiral model of pitch: Interrelations with musical and psychoacoustic scales and with cochlear parameters. James D. Miller (Dept. of Speech and Hearing Sci., Indiana Univ., 200 S. Jordan Ave., Bloomington, IN 47405, jamdmill@Indiana.edu)

A spiral model of pitch can represent pitch chroma and pitch height. Its arclength provides the frequency-position map of the cochlea, and, when scaled by the number of auditory-ganglion cells, provides the Mel scale. Pitch-like scales fall into three groups. Group I is typified by the Cent scale, which can be expressed as a pitch spiral. Other Group-I scales are musical notation, logarithmic scales of frequency, Fechners Law for complex tones, and the pitch helix. Group II is typified by Greenwoods cochlear map, which is simply related to the arclength of the pitch spiral. Other Group-II scales are the sums of the number of (a) equivalent-rectangular bandwidths (NERBs); (b) difference limens for frequency (NDFs); and (c) inner-hair cells (NIHCs) and are consistent with critical ratios. Group III is typified by the Mel Scale, which can be derived from the pitch spiral and the distribution of auditory-ganglion cells. Other Group-III scales are the sums of the number of (a) auditory ganglion cells (NSGCs), (b) frequency-modulation-difference limens (NFMDLs), and (c) Barks (NBARKs). The pitch spiral seems a basic concept from which other pitch-like scales can be derived.

4aPP22. Repetition “pitch” in chinchillas. William P. Shofner and William M. Whitmer (Parmlly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, wshofne@luc.edu)

Repetition pitches are evoked in human listeners by rippled noises. Infinitely iterated rippled noise (IIRN) is generated when wideband noise is delayed, attenuated, and added to the original wideband noise through positive (+) or negative (–) feedback. In human listeners, the pitch of IIRN(+) is matched to the reciprocal of the delay, whereas the pitch of IIRN(–) is an octave lower. A stimulus generalization paradigm was used to characterize IIRN repetition pitch in chinchillas. Chinchillas were trained to discriminate IIRN(+) with a 4-ms delay from IIRN(+) with a 2-ms delay. In the generalization task, chinchillas were tested with IIRN(+) having delays between 2–4 ms. The pitches evoked by these IIRNs ranged from 500–250 Hz. The delayed noise attenuation for all IIRNs was fixed at –1 dB. A systematic gradient in behavioral response occurred along the dimension of delay for each animal tested, suggesting that a perceptual dimension of pitch exists. Responses to IIRN(–) evoking pitches between 500–250 Hz (delays of 1–2 ms) were also measured. Responses to IIRN(–) were more variable among animals, suggesting that other perceptual cues such as timbre differences may be stronger than the pitch cues. [Work supported by NIDCD R01 DC005596.]

4aPP23. Influence of beats on mistuning detection in consonant and dissonant musical intervals. Craig E. Lewiston and Andrew J. Oxenham (MIT Res. Lab. of Electron. and Harvard-MIT Speech & Hearing Bioscience & Technol. Prog., 77 Massachusetts Ave., Cambridge, MA 02139, lewiston@mit.edu)

This study investigated the mechanisms underlying the detection of mistuning in musical intervals. Musical consonance is thought to be defined in part by an absence of cochlea-generated beats. Dissonant intervals, and mistuned consonant intervals, produce a percept of beats. Thus, mistuning of consonant intervals might be detected through the presence of beats, whereas mistuning of dissonant intervals might be detected by a change in the pattern of beats. Alternatively, mistuning might be detected by a direct comparison of the two pitches in each interval. In our experiment the presence of beats was controlled by presenting the two tones to both ears or to separate ears, by using either pure tones or complex tones, and by presenting the tones sequentially or concurrently. Results using highly trained musicians showed large threshold elevations (poorer performance) in the absence of cochlea-generated beats, suggesting that the mistuning of concurrent sounds in normal circumstances is mediated by beat detection. However, even in the absence of beats, mistuning detection

thresholds for dissonant intervals, such as the tritone or minor second, were poorer than for consonant intervals, such as the octave or fifth. [Work supported by NIH grants T32DC00038 and R01DC05216.]

4aPP24. Further explorations of the contribution of a nonsimultaneous mistuned harmonic to residue pitch. Hedwig E. Gockel, Robert P. Carlyon (MRC Cognition and Brain Sci. Unit, 15 Chaucer Rd., Cambridge CB2 2EF, UK, hedwig.gockel@mrc-cbu.cam.ac.uk), and Christopher J. Plack (Univ. of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK)

Ciocca and Darwin [V. Ciocca and C.J. Darwin, *J. Acoust. Soc. Am.* **105**, 2421–2430 (1999)] reported a surprising finding: The shift in residue pitch caused by mistuning a harmonic was the same when the mistuned harmonic was presented after the remainder of the complex as when it was

simultaneous. The present study tried to replicate this result, and investigated the role of the presence of the nominally mistuned harmonic in the matching sound. Subjects adjusted a matching sound so that its pitch equaled that of a subsequent 90-ms complex tone (12 harmonics of a 155-Hz F_0), whose mistuned ($\pm 3\%$) 3rd harmonic was presented either simultaneously with or after the remainder. In experiment 1, the matching sound was a harmonic complex whose 3rd harmonic was either present or absent. In experiment 2, it was a sinusoid. In experiment 3, the target and matching sound contained non-overlapping harmonics. In all experiments, a non-simultaneous mistuned component produced significantly smaller pitch shifts than a simultaneous one. In the absence of the nominally mistuned harmonic in the matching sound, the pitch shift with simultaneous presentation was about five times larger than that with non-simultaneous presentation of the mistuned component. [Work supported by EPSRC Grant GR/R65794/01.]

THURSDAY MORNING, 19 MAY 2005

PLAZA B, 9:00 TO 11:45 A.M.

Session 4aSA

Structural Acoustics and Vibration and Musical Acoustics: Vibration of and Acoustic Radiation from Musical Instruments I

Courtney B. Burroughs, Cochair

Applied Research Lab., Pennsylvania State Univ., P.O. Box 30, State College, PA 16804-0030

Thomas D. Rossing, Cochair

Physics Dept., Northern Illinois Univ., De Kalb, IL 60115

Invited Papers

9:00

4aSA1. The banjo in the time-domain. Joseph Dickey (The Whiting School of Eng., Johns Hopkins Univ., 810 Wyman Park Dr. Ste. G010, Baltimore, MD 21211)

The banjo is under-represented in the musical acoustics literature. This is the case even though it is basically a vibrating and radiating membrane driven by a plucked string, and thereby lends itself naturally to modeling and analysis. A recent frequency domain model, [Dickey, *J. Acoust. Soc. Am.* **114**, 2958–2966 (2003)] will be reviewed. The work reported here is a time domain analysis. First, the response of an isolated, plucked, lossy and dispersive string is presented. The string is then connected through a bridge to the center of a circular membrane. The string and head are treated as two fully connected dynamic systems and the response of the head is calculated. The isolated head is then driven off-center with an harmonic drive and the temporal response is determined by ray-tracing and summing amplitudes over a grid. Resonance frequencies and modal patterns agree well with experiment thereby validating the technique. Finally, the resonant string is considered as an external, transient, off-center and spatially distributed drive and the temporal response of the head is derived using the ray tracing technique. These modal patterns evolve in time and eventually stabilize to patterns that are in qualitative agreement with experiment.

9:30

4aSA2. Structural vibrations and sound radiation fields of classical guitars. Bernard E. Richardson (School of Phys. and Astron., Cardiff Univ., 5 The Parade, Cardiff CF24 3YB, UK, RichardsonBE@cardiff.ac.uk)

In its few lines, Federico Garcia Lorca's poem "Las Seis Cuerdas" (the six strings) invokes a beautiful image of sound production on the guitar. It is the player's interaction with the six strings which makes the music, but were it not for the body of the guitar, the instrument would have no voice. Most studies of the acoustical function of guitars have concentrated on measuring the structural vibrations or the sound radiation fields in isolation of the strings. These studies fail to encapsulate the important aspects of string-body coupling, which has a marked influence on the decay rates of string vibrations, and the radiativity of individual body modes. By measuring both input admittance at the bridge and sound-field topology for individual modes of the guitar, it is possible, through psychoacoustical experiments, to investigate the relative importance of the various acoustical parameters which define the response of the instrument. This paper will discuss experimental techniques and present data on ten classical guitars which go some way to identifying key components in the low- to mid-frequency mechanical and acoustical actions of the guitar. [Work originally supported by the Leverhulme Trust.]

10:15

4aSA3. Development of a composite material concert harp soundboard to match the structural acoustic performance of a wooden soundboard. Thomas J. Royston, Melinda J. Carney, Curt Preissner, and John Roxworthy (Acoust. & Vib. Lab., Univ. of Illinois at Chicago, 842 W. Taylor St., MC 251, Chicago, IL 60607, troyston@uic.edu)

The replacement of a Sitka spruce grand concert harp soundboard with a carbon fiber-reinforced plastic soundboard could provide improved durability and long-term stability, helping to mitigate stresses relating to fluctuating temperature and humidity that are inherent in wooden instruments. This presentation outlines the development of a composite soundboard by a combined experimental and computational approach. Experimental modal analysis data was compared with the results from a computational finite-element model to find the effective material properties of the multi-layered laminate wooden soundboard, treating the veneer, soundboard and varnish as one material. With the effective wood material properties determined, a composite test section was designed and fabricated based on specific matching criteria, duplicating the behavior of the wood with the composite laminate. The experimental natural frequencies and mode order of the composite test section closely matched the experimental results of the wood section within a 10% difference. Given these results, the method was then applied to design and manufacture a full composite soundboard. Various finite element models were created to develop the most practical design while sufficiently matching the wood soundboard properties. Issues regarding the implementation of the composite soundboard were also investigated.

10:45

4aSA4. Modes of vibration and sound radiation from percussion instruments. Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115, rossing@physics.niu.edu)

When a membrane or bar or plate is struck, it vibrates in a complex manner, which can be described in terms of normal modes of vibration. We describe the modes of vibration of percussion instruments, such as drums and bells, some ways in which the normal modes are observed, and the way in which they determine sound radiation from the instruments.

11:15

4aSA5. Interferometric studies of a piano soundboard. Thomas R. Moore (Dept. of Phys., Rollins College, Winter Park, FL 32789, tmoore@rollins.edu)

Ongoing efforts to understand and model the dynamics of the modern piano are hampered by a lack of understanding of the deflection shapes of the soundboard, which are extremely complicated due to the complex construction. Often deflection shapes of harmonically vibrating objects can be determined using holographic or speckle pattern interferometry, but in practice the stability necessary to implement these methods is difficult to achieve for a large wooden structure such as a piano. We show theoretically and experimentally that the deflection shapes of large objects that are typically too unstable for interferometry can be determined by modifying the common form of the electronic speckle pattern interferometer. Furthermore, using this modified interferometer the decorrelation of the speckle due to ambient vibrations actually enhances the precision of the interferogram. We discuss some interesting observations of the deflection shapes of the soundboard of a fully assembled piano, compare the deflection shapes of the lowest resonances with simple models, and demonstrate how the interferograms can be used to determine the driving point impedance.

Session 4aSC

Speech Communication: Advanced Methods in Speech Research and Speech Technology (Poster Session)

Terrance M. Nearey, Chair

*Dept. of Linguistics, Univ. of Alberta, 4-32 Assiniboia Hall, Edmonton, AB T6G 2E7, Canada**Contributed Papers*

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon.

4aSC1. Development of an anthropomorphic talking robot and the mimicking speech control. Kotaro Fukui, Kazufumi Nishikawa, Toshiharu Kuwae, Atsuo Takanishi (Dept. of Mech. Eng., Waseda Univ, 3-4-1 Ookubo, Shinjuku-ku, Tokyo, Japan), Hideaki Takanobu (Kogakuin Univ., Tokyo, Japan), Takemi Mochida (Commun. Sci. Labs., NTT, Kanagawa, Japan), and Masaaki Honda (Waseda Univ., Saitama, Japan)

We developed an anthropomorphic talking robot WT-4 (Waseda Talker No. 4) to produce human speech. WT-4 consists of 1-DOF lungs, 4-DOF vocal cords and articulators (the 7-DOF tongue, 5-DOF lips, 1-DOF teeth, nasal cavity and 1-DOF soft palate); the total DOF is 19. The lips and the tongue are made of elastic material to allow large deformation and to prevent from the air and sound leaks, and are controlled by a looped wire mechanism to form their various configurations. The talking robot enables to produce vowel and consonant sounds by mimicking the vocal cords vibration and the fricative and plosive source generation by the air flow as well as dynamically controlled vocal tract acoustic resonance in human speech production. Articulatory control of the talking robot is designed to track the acoustic goals (pitch, sound power, two formant frequencies, and voice-unvoiced timing) of the speech. The robot parameters are determined by minimizing the weighted mean squared error of these acoustic parameters between the human and robot speech sounds. It is shown that this mimicking speech control is effective in producing fluent continuous speech by the talking robot.

4aSC2. The tube resonance model speech synthesizer. Leonard C. Manzara (Dept. of Comput. Sci., Univ. of Calgary, 2500 Univ. Dr. NW, Calgary, AB, Canada T2N 1N4, manzara@cpsc.ucalgary.ca)

The Tube Resonance Model (TRM) synthesizer is an articulatory speech synthesizer implemented in software. It directly emulates the resonant behavior of the oropharyngeal and nasal tracts using digital waveguides. The oropharyngeal cavity is subdivided into 8 regions of unequal length, where particular regions correspond to the human articulators of tongue, teeth, and mouth. The radius (cross-sectional area) of each region can be varied independently over time. The differences in radii between regions gives rise to differences in acoustic impedance, which is modeled using two-way scattering junctions. The nasal cavity is composed of 5 equal-length sections, and is connected to the vocal tract via another section (the velum) using a three-way scattering junction. The total length of the tube can be varied over a continuous range, allowing one to synthesize male, female, and juvenile voices.

4aSC3. Study of effect of speaker variability and driving conditions on the performance of an automatic speech recognition engine inside a vehicle. Shubha Kadambe (HRL Labs., LLC, 3011 Malibu Canyon Rd., Malibu, CA 90265)

Spoken dialogue based information retrieval systems are being used inside vehicles. The user satisfaction of using such a system depends on how an ASR engine performs. However, the performance of an ASR is affected by speaker variability, driving conditions, etc. Here, we report the study that we performed to analyze these effects of speaker variability, different driving conditions and the effect of driving task on the ASR performance. This study consists of experimental design, data collection and systematically testing an ASR engine using this data. From the obtained results, it can be observed that (I) the ASR performance exhibits (a) significant speaker variability since the stress of driving task varies from speaker to speaker, (b) significant performance degradation across driving conditions since the noise type and level varies and (c) significant effect of driving task on recognition performance, and (II) the effect of live noise on recognition performance is not same as adding car noise to the pre-recorded speech data. The former observation is important since by just training an ASR engine on lots of speech data will not help and it is essential to include stress factors and cognition load in ASR engines to improve its performance.

4aSC4. A noise-reduction strategy for speech based on phase-opponency detectors. Om Deshmukh (Dept. of Elec. and Comp. Engr. and Inst. for Systems Res., Univ. of Maryland, College Park, MD 20742), Michael C. Anzalone (Syracuse Univ., Syracuse, NY 13224), Carol Y. Espy-Wilson (Univ. of Maryland, College Park, MD 20742), and Laurel H. Carney (Syracuse Univ., Syracuse, NY 13224)

A noise-reduction algorithm was developed based on a neural model for detection that is robust in fluctuating noise. The phase-opponency (PO) neural model correlates the outputs of two different auditory filters that differ in phase by 180 degrees at the target frequency. The PO detector used here consists of a pair of overlapping bandpass filters with phase responses that differ by 180 degrees near the center frequency (CF). The correlation between the filter responses is reduced when a narrowband signal near CF is present in a noisy background. A bank of PO detectors was used to process speech corrupted by additive Gaussian noise. The time-varying outputs of the detectors can be post-processed to retain information-rich regions, such as formants and frication onsets, while greatly reducing noise between formants. The final detector output controlled the gains in a separate analysis/synthesis filterbank. Spectrograms of speech sounds before and after noise reduction illustrate the ability of the system to detect major features in speech down to low signal-to-noise ratios. The quality of the processed signal will be demonstrated. This system is intended as a front-end for speech recognition systems or in

4aSC5. A new method of extracting the filter characteristics of the nasal cavity using homorganic nasal-stop sequences. Hansang Park (Dept. of English Education, Hongik Univ., 72-1 Sangsu-dong, Mapo-gu, Seoul, Korea)

This study attempts to derive the filter characteristics of the nasal cavity of individual speakers. Since the only difference between a nasal and a homorganic voiced stop, such as [mb] and [nd], is whether the passage to the nasal cavity is open or not, the subtraction of the LPC spectrum of the voiced stop from that of the preceding nasal leads to the filter characteristics of the nasal cavity of an individual speaker regardless of place of articulation. The results showed that the spectral differences between samples of 20 ms taken from the steady states of the nasal and the following voiced stop were close to constant regardless of place of articulation, representing characteristic poles and zeroes, and that the spectral differences varied with speakers. This study is significant in that it provides a new method of extracting the filter characteristics of the nasal cavity, and that the spectral difference between a nasal and a homorganic voiced stop can be used as a parameter of the filter characteristics of the nasal cavity of individual speakers.

4aSC6. Synthesizing speech acoustics from head and face motion. Adriano V. Barbosa, Hani C. Yehia (CEFALA/PPGEE, Universidade Federal de Minas Gerais, Av. Antonio Carlos, 6627, Belo Horizonte, MG, 31270-010, Brazil, adriano.vilela@gmx.net), Andreas Daffertshofer (Vrije Universiteit, Amsterdam, The Netherlands), and Eric Vatikiotis-Bateson (Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z1)

This work outlines a quantitative analysis of the relation between speech acoustics and the face and head motions that occur simultaneously [A. V. Barbosa, Ph.D. thesis, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, 2004]. 2-D motion data is obtained by means of a video camera. An algorithm has been developed for tracking markers on the speaker's face from the acquired video sequence [A. V. Barbosa, E. Vatikiotis-Bateson, and A. Daffertshofer, in Proceedings of the 8th ICSLP Interspeech 2004, Korea, 2004]. The motion domain is represented by the 2-D marker trajectories, whereas line spectrum pairs (LSP) coefficients and the fundamental frequency F_0 are used to represent the speech acoustics domain. Mathematical models are trained to estimate the acoustic parameters (LSPs + F_0) from the motion parameters (2-D marker positions). The estimated acoustic parameters are then used to synthesize the acoustic speech signal. Cross-domain analysis for undecomposed (i.e., full head + face) and decomposed (i.e., separated head and face) normalized 2-D motions is performed. Syntheses from each method using intelligibility tests and qualitative comparison of the original and synthesized utterances are being evaluated.

4aSC7. ArtiSynth designing a modular 3D articulatory speech synthesizer. Florian Vogt, Oliver Guenther, Allan Hannam, Kees van den Doel (Univ. of British Columbia, 2356 Main Mall, Vancouver, BC, Canada V6T 1Z4, fvogt@ece.ubc.ca), John Lloyd, Leah Vilhan, Rahul Chander, Justin Lam, Charles Wilson, Kalev Tait, Donald Derrick, Ian Wilson, Carol Jaeger, Bryan Gick, Eric Vatikiotis-Bateson, and Sidney Fels (Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z4)

ArtiSynth is a modular, component-based system for performing dynamic 3D simulations of the human vocal tract and face. It provides a test bed for research in areas such as speech synthesis, linguistics, medicine, and dentistry. ArtiSynth's framework enables researchers to construct, refine, and exchange models of all parts of the vocal tract and surrounding structures. ArtiSynth introduces a probe concept to unify input and output data flow, which allows control of and access to models with time varying data series. ArtiSynth supports interconnected heterogeneous models, such

as rigid body, mass-spring, and parametric, using a point-set connection method, called markers, for constraint satisfaction. Using ArtiSynth, we created a muscle-driven rigid body jaw model, a parametric principle component tongue model from MRI images, a parametric lip model, and mass-spring face tissue model. We combined them in various ways. Data from medical imaging (MRI, CT, and ultrasound) and other technologies such as optical tracking can be used to drive ArtiSynth models. We are currently developing an acoustical rendering framework supporting source-filter models and other advanced methods. The system incorporates a powerful scripting interface as well as an easy-to-use graphical interface. [Work supported by NSERC Canada and ATR Japan.]

4aSC8. Design of a 6 degree of freedom anthropomorphic robotic jaw. Edgar Flores and Sidney Fels (Dept. of Elec. & Comput. Eng., UBC, 2356 Main Mall, Vancouver, BC, Canada V6T 1Z4)

We have created a 6 DOF robotic jaw capable of producing, in real-time, the complex set of motions described by the human jaw during speech or mastication. The jaw is designed to fit within a larger robotic human figure such as the head, neck and torso of the 25 DOF Infanoid. [Kozima, Hideki: Infanoid: A Babybot that Explores the Social Environment, K. Dautenhahn *et al.* (eds.), Socially Intelligent Agents: Creating Relationships with Computers and Robots, Kluwer Academic Publishers, pp. 157–164, 2002]. The produced mechanical prototype has been designed to accommodate a prosthesis mandible with dentures. The mechanism could fit within the skull of the average man; where it would occupy less than 1/3 of the skull cavity. Two TMJs (temporomandibular joints) support the prosthesis, where each is driven by a 3 DOF parallel manipulator. In order to combine the motion of both manipulators each TMJ is capable of 3 DOF. The system is controlled via a USB port using software that models the human skull including collision detection mechanisms. The jaw allows for linear control, zero-backlash, and up to three times exaggerated mobility ranges making it also suitable for speech research, facial gesture affect research and dentistry applications.

4aSC9. Effects of subglottal acoustics on phonation onset. Juergen Neubauer, Zhaoyan Zhang, and David Berry (UCLA School of Medicine, 31-24 Rehabilitation Ctr., 1000 Veteran Ave., Los Angeles, CA 90095, zyzhang@ucla.edu)

The effect of subglottal acoustic loading on the vocal fold vibration was investigated using a self-oscillating mechanical model of the folds. Although the influence of the supraglottal tract on vocal fold vibration has received more attention than the subglottal system, the influence of the subglottal system on vocal fold vibration is also potentially significant, and merits investigation. In this study, the subglottal system consisted of a uniform tube connected to an expansion chamber on the flow supply end (e.g., a pseudo-lung). The length of the subglottal tube was varied systematically over a relatively large range in order to investigate the influence of subglottal acoustics on vocal fold vibration. Phonation onset and offset pressures were measured in the subglottal tube as a function of tube length. Over the range of investigation, the fundamental frequency of phonation was found to be negatively correlated with the subglottal tube length. However, both phonation onset and offset pressure were positively correlated with subglottal tube length, with the onset pressure increasing faster than the offset pressure. This hysteresis effect vanished and the two pressures merged at a small value of the subglottal tube length, indicating a change in the onset behavior from a subcritical Hopf bifurcation to a supercritical Hopf bifurcation (a codimension-2 bifurcation point). In addition, phonation did not exist below a critical subglottal tube length.

4aSC10. Classification of initial segments of sounds using onset data structures. Dagmar S. Fraser and Leslie S. Smith (Dept. of Computing Sci. and Mathematics, Univ. of Stirling, Stirling FK9 4LA, Scotland, dsf@cs.stir.ac.uk)

Biologically motivated techniques are used to record onset data structures [L. S. Smith and D. S. Fraser, IEEE TNNS **15**, 1125–1134 (2004)]. These structures record sets of events detailing both the spectral and intensity structure of sound onsets. They are inspired by the output spikes from onset cells in the cochlear nucleus. We suggest that the spectro-temporal characteristic of these structures provides useful information for classifying initial segments of sounds. The original onset data structure is of high dimensionality, and is difficult to interpret directly. A self-organizing feature map (SOFM or Kohonen network) is therefore used to provide a lower dimensional coding. The SOFM requires fixed-length data vectors, preferably not of too high a dimensionality. A number of different reformatting techniques are used to provide this. For the SOFM, identified activated map areas are associated with particular sounds. Using the TIMIT database these areas can be labelled, providing a testable classification scheme. This is compared with other work classifying TIMIT phonemes. We believe that the SOFM can be extended to classify the onsets of other (non-speech) sounds. [Work supported by UK EPSRC.]

4aSC11. Computational fluid dynamics simulation of a sibilant. Kazunori Nozaki, Toyokazu Akiyama, Shinji Shimojo (Cybermedia Ctr., Osaka Univ., 5-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan, nozaki@cmc.osaka-u.ac.jp), Shingo Maeda, Takeshi Kaishima (Osaka Univ.), Hiroo Tamagawa, and Yoshinobu Maeda (Osaka Univ.)

It is well known that a sibilant is generated around the anterior teeth and the frontal part of palate. However, there is no study that the sibilant sound source is detected and visualized. Dentists often need to fabricate the prostheses including the anterior teeth and the palate. It would be useful if the critical part of the prostheses on the generation of sibilant could be detected before the fabrication and treatment. In this present study, Computational Fluid Dynamics simulation was performed in order to examine the oral air flow on pronouncing sibilant. Moreover, Powell sound source around the anterior teeth was detected and visualized by combining Computational Fluid Dynamics simulation and Computational Acoustical Analysis. As a result of this study, dentists can obtain the clinical criteria of the prosthetic treatment with considering sibilant.

4aSC12. Deriving speaking rate effects on tonal realization without varying the speech rate. Alan C. L. Yu (Univ. of Chicago, 1010 E 59th St., Chicago, IL 60637, aclyu@uchicago.edu)

Previous research investigates the effect of durational variation on tonal realization by artificially inducing a change in speaking rate. Through an investigation of the f_0 pattern of Cantonese mid-rising tone, this paper introduces a novel, indirect, method of investigating durational effect on tonal realization by looking at the variation in tonal realization in different syllable types. Cantonese has three syllable types: CV, CV(V)O, CV(V)N. CV(V)N syllables are found to be significantly longer than CV syllables, which in turn are significantly longer than CV(V)O syllables. In this study, five native Cantonese speakers were asked to recite a list of Cantonese words with a mid-rising tone in a carrier phrase. The f_0 of the Turning Point (the lowest point of a rise), the f_0 peak of the rise, the durations of the Turning Point and the f_0 peak relative to the onset of voicing, and the duration of the syllable rime were measured. The results show that the f_0 peak frequency remains constant regardless of the duration of the syllable. The f_0 of the Turning Point gets higher when the duration of the syllable becomes shorter. Finally, the slope of the f_0 rise remains constant regardless of the duration of the syllable.

4aSC13. Mechanical vocal cord model mimicking human biological structure. Eiji Shintaku, Kotaro Fukui, Kazufumi Nishikawa, Shunsuke Ikeo (Dept. of Mech. Eng., Waseda Univ., 3-4-1 Ookubo, Shinjuku-ku, Tokyo 169-8555, Japan), Kentaro Takada, Atsuo Takanishi (Waseda Univ., Saitama, Japan), Hideaki Takanobu (Kogakuin Univ., Tokyo, Japan), and Masaaki Honda (Waseda Univ., Saitama, Japan)

We present a mechanical vocal cord model aiming for a talking robot, WT-5 (Waseda Talker No. 5). Unlike a musical reed which has been used in conventional mechanical speech synthesizer, the vocal cord model is formed to mimic the human's vocal cord in the shape and the biological structure. It is made of a thermoplastic rubber, Septonh (Kuraray Co. Ltd.) of which the elasticity like a human's, and has 3-DOF mechanisms which is similar to the human structure. 1-DOF link mechanism could change the pitch by stretching the length of the vocal cords. The 2-DOF arm mechanism is used to mimic the abduction and adduction of a human arytenoid cartilage. The vocal cord model was excited by air flow exhausted from a mechanical lung model. The vibration pattern was observed by a high-speed camera, and the glottal volume velocity and the sound pressure were recorded by a mask-type wire screen pneumotachograph and a microphone. It was shown that the lower and upper edges of the vocal cords could vibrate in a different phase and the sound spectrum was similar to the source spectrum of human speech.

4aSC14. An educational articulatory synthesizer, EASY. Richard S. McGowan (CReSS LLC, 1 Seaborn Pl., Lexington, MA 02420) and Reiner Wilhelms-Tricarico (CReSS LLC and Haskins Labs., New Haven, CT 06511)

An articulatory synthesizer has been written for general educational and research use. It has been named Educational Articulatory Synthesizer, EASY, and written in the MATLAB programming language. In its current instantiation, the synthesizer performs calculations in the frequency domain and includes the effects of wall-vibration, viscous and thermal boundary layer loss, and radiation loss. Sources anywhere from the glottis to the mouth can be specified. Further, the code allows for side branches other than the nasal tract so that sounds such as laterals can be specified. A low-frequency aerodynamic module has also been included. Midsagittal shape is controlled by a hierarchy of flesh points. Points high in the hierarchy specify the overall shape by determining reference positions for points lower in the hierarchy, which, in turn, determine the shape locally. [Work supported by grant NIDCD-001247 to CReSS LLC.]

4aSC15. The relation between learning Mandarin Pinyin or Zhuyin and L2 (English) production. Yan Helen Yu and Fredericka Bell-Berti (Dept. of Speech, Commun. Sci., & Theatre, St. John's Univ., 8000 Utopia Pkwy, Jamaica, NY 11439)

This study investigates the relation between early written first language learning experience and the perception and production of speech in a person's second language (L_2). It compares the English pronunciation of native Mandarin speakers who have had different sound-annotating learning experiences (Pinyin versus Zhuyin). We predict that native Mandarin speakers who used the Pinyin system will produce some English sounds as their Pinyin counterparts if the same Roman letter is used as a pronunciation symbol both in English and in Pinyin system. However, because Pinyin and Zhuyin are not the primary written form for Mandarin, and are extensively used only in the first years of school, and because L_2 learning involves exposure to speech as well as reading, we expect only moderate differences between the two groups' perception and production abilities. We will examine the interaction for Pinyin and Zhuyin learners between (Mandarin) L_1 learning experience, age effects, L_2 experiences, and the establishment of English phonetic categories for the English monophthongal vowels and consonants /r/, /z/, and /l/.

4aSC16. Characterization of voice pathologies from acoustic signals using wavelet analysis. Sarah A. Bentil and Yuling Yan (Dept. of Mech. Eng., 302 Holmes Hall 302, Univ. of Hawaii-Manoa, Honolulu, HI 96822, bentil@hawaii.edu)

We present a method to quantitatively characterize voice abnormalities using wavelet analysis. The proposed method uses wavelets to decompose acoustic signals, acquired clinically from patients with normal and pathological voices, into their designated frequency/time components. These components contain valuable information on the unique dynamic properties of the vocal system and correlate with specific voice conditions. A comparative analysis of these vocal signals using spectrogram is also presented. These combined analyses provide comprehensive representation and quantitative characteristics of the vocal dynamics, which may provide key indication of vocal abnormalities. Further wavelet analysis of acoustic data evaluates variations in the characteristics of the vocal signal from one glottal cycle to the next. This is valuable since vocal signals representing most pathological voice productions exhibit inter-cycle variations in intensity or/and frequency. Our results of analysis show that the wavelet analysis can be used to characterize voice pathologies and provide information regarding the type and severity of the disorder. [Work supported by NSF awarded to Yan.]

4aSC17. Synthesizer software for modeling voice quality. Norma Antonanzas-Barroso, Bruce R. Gerratt, and Jody Kreiman (Div. of Head/Neck Surgery, UCLA School of Medicine, 31-24 Rehab Ctr., Los Angeles, CA 90095-1794, jkreiman@ucla.edu)

This poster presents a formant synthesizer that is designed especially for detailed modeling of voice quality. Users may interactively manipulate source pulse shapes and/or spectral characteristics of the voice source, the noise spectrum and noise-to-signal ratio, jitter and shimmer, vocal tremor rate and extent, and the vocal tract transfer function (formants and bandwidths), all in near-real time. Pitch contours can be modeled directly or imported. The synthesizer may be utilized in method-of-adjustment tasks, or sequences of stimuli may be created for use in other experimental paradigms. An interactive audio-visual display makes this a useful tool for teaching voice acoustics as well. Both an executable version of the program and the underlying code will be available during the conference. Copies of software that supports the voice modeling effort, including an interactive inverse filter, will also be distributed. [Work supported by NIDCD grant DC01797.]

4aSC18. Simulation and analysis of tremor in speech production. Kimberly A. Farinella and Brad H. Story (Dept. of Speech, Lang., and Hearing Sci., Univ. of Arizona, P.O. Box 210071, Tucson, AZ 85721-0071)

Tremor of the muscles used for voice and speech production creates frequency and amplitude modulations in the acoustic speech signal. Frequency modulation results primarily from changes in muscle activations that alter the mass and stiffness of the vocal folds, consequently altering the fundamental frequency (F_0). The primary source of amplitude modulation comes from time-varying activation of the respiratory muscles (which alter alveolar pressure), from laryngeal muscles that modify the maximum glottal area, or from interaction of voice source harmonics with the vocal tract filter. The purpose of this preliminary investigation was to analyze acoustic signals generated with an articulatory speech synthesizer in which tremor was imposed separately at each of the three anatomical sites: the respiratory system, larynx, and upper airway (pharynx). The synthesizer consisted of a voice source model of the time-varying glottal area coupled to a wave propagation model of the airways upstream and downstream of the vocal folds. This allowed for complete control of F_0 , alveolar pressure, glottal area, vocal tract shape, and modulation frequencies of each. Frequency and amplitude contours of the resulting acoustic

signals were extracted from the waveform and analyzed and compared using customized Matlab routines. Submitted For (Speech Communication) Young Presenter Award and (Speech Communication) Best Student Paper Award.

4aSC19. Temporal features in TV news and weather forecasts. Tatiana I. Shevchenko and Natalia Uglova (Moscow State Linguistic Univ., 38 Ostozhenka, Moscow 119992, Russia)

The paper is aimed at investigating the specific features of temporal component as manifested in authentic TV speech, news and weather forecasts of the three channels: NBC, Texas News, Philadelphia Fox. In our 10 min corpus (9 speakers, 5 men and 4 women) durational patterns and pauses between them as well as their ratio were measured. The data showed that on average a durational pattern lasts 2655 ms, while an average pause duration is 193 ms. Thus pausing does not take a major amount of time, while phonation is 14 times as long. Furthermore, the analysis reveals the abundance of short and extra short pauses which also stands for a rapid tempo. The temporal characteristics were correlated with regional affiliation, genre and gender distinctions. The speakers demonstrated uniformity in the basic temporal features determined by time limit constraints. The results of our investigation basically agree with previous research based on mass media interviews, reading and spontaneous monologues. However, our findings show that a salient feature of newsreaders speech is their ability to deliver an enormous volume of information at a very quick tempo, which is a real challenge to viewers.

4aSC20. Stabilization techniques for ultrasound imaging of speech articulations. Lisa Davidson and Paul De Decker (Dept. of Linguist., New York Univ., 719 Broadway 4th Fl, New York, NY 10003)

One challenge for the ultrasound imaging of the tongue during connected speech is stabilization of both speakers heads and the ultrasound transducer. To accurately analyze tongue shapes, researchers must ensure that differences among images result from changes in tongue shape, not head or transducer movement. In this validation study, we present an inexpensive, space-saving technique for stabilizing the head and transducer during the collection of ultrasound images in the laboratory. Four speakers were video-recorded producing 5 blocks of 19 sentences each. The speakers' heads were immobilized with a moldable head stabilizer (Comfort Company, model #HSM) affixed to a wall while the transducer was held by a microphone stand. Markers were attached to points on the speakers' faces and the transducer. Custom-written MATLAB software was used to find the center of the markers for every 10th frame of the recordings. Results indicate that the transducer does not move, and head movement beyond the measurement error of 1 mm is confined to the first block, while speakers are adjusting to the equipment. As compared to other stabilization techniques [Stone and Davis, J. Acoust. Soc. Am. **98**, 3107–3112 (1995)], this method provides equivalent immobilization that is significantly less expensive and portable for fieldwork.

4aSC21. Acoustic eigenmodes and formant-cavity affiliations for the time-varying vocal tract. Gordon Ramsay (Haskins Labs., 270 Crown St., New Haven, CT 06511-6204, ramsay@haskins.yale.edu), Philip Rubin (Haskins Labs., New Haven, CT 06511-6204), and Catherine Best (Haskins Labs., New Haven, CT 06511-6204)

Acoustic analyses of speech production are often based on assumptions about formant-cavity affiliations. Typically, these are derived from approximations based on coupled resonators, or from small-perturbation analyses for open vowel configurations. Few studies have accurately calculated formant-cavity affiliations for tightly-constricted vocal tract shapes, or for vocal tract shapes that change in time. In this paper, we show how to determine the acoustic eigenmodes of the vocal tract from a physical simulation, and apply this to examine time-varying formant-cavity affiliations during sequences of consonants and vowels. A state-

space model of quasi-one-dimensional acoustic wave propagation in a time-varying elastic tube is derived, using the finite volume method. The eigenvalues and eigenvectors of the resulting implicit matrix recursion are shown to determine the formant frequencies and bandwidths, and the spatial distribution of potential and kinetic energy for each formant, at each

point in time. Illustrations are given for examples of VCV sequences constructed using an articulatory model, including stops, fricatives, and approximants in different vocalic contexts. Formant-cavity affiliations, as defined by the localization of acoustic energy along the vocal tract, do not necessarily correspond to actual cavities. [Work supported by NIH.]

THURSDAY MORNING, 19 MAY 2005

GEORGIA A, 8:30 TO 11:45 A.M.

Session 4aSP

Signal Processing in Acoustics: Design and Applications of Arrays

Paul J. Gendron, Chair

Naval Research Lab., Acoustics Div., 4555 Overlook Ave., SW, Washington, DC 20375-0002

Contributed Papers

8:30

4aSP1. Broadband nearfield beamformer with compact size by using Nash genetic algorithm. Soonkwon Paik (Elect. & Comput. Eng., Univ. of Texas, 1 Univ. Station Stop C0803, Austin, TX 78712, soonkwon@mail.utexas.edu) and Elmer L. Hixson (Univ. of Texas, Austin, TX 78712)

Though the acoustical array is effective in the reverberant field, its usage has been limited at low frequencies since the array with the length of more than 1 meter is difficult to install inside a vehicle. Furthermore, it is impractical to use circular coordinates especially inside vehicle since the sound source, a human mouth, stays at a fixed distance from and moves along the parallel line to the surface of the headliner. For the delay-and-sum beamformer in near-field, it is difficult to get the complex coefficients, amplitude weight and time delay, by analytical methods. In this paper, one numerical optimization method for a fixed microphone array is investigated which uses the Nash Genetic Algorithm (Nash GA) which was originally used in electromagnetism problems. By using Nash GA, first the optimal geometry of array element alignment inside a vehicle is found, then the complex coefficients are obtained. The results measured in an anechoic chamber correspond to that of computer simulation and satisfy the requirement of hands-free mobile telephony where the microphone array is installed in a vehicle's headliner and the data is also measured in a reverberant field to investigate the signal to noise ratio.

8:45

4aSP2. Capacity of the oceanic waveguide. W. J. Higley, Philippe Roux, and W. A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., 9500 Gilman Dr., La Jolla, CA 92093-0238)

Capacity between two multiple-input/multiple-output (MIMO) arrays can be computed by decomposing the Green's function matrix into orthogonal functions, which can be thought of as independent subchannels. In a waveguide with sufficient array sampling, the modes are the independent subchannels, and capacity can easily be calculated for given waveguide parameters. Capacity is calculated using the optimal power allocation strategy of "water-filling." Water-filling is simultaneously done across modes and frequency. Theoretically computed capacities for an oceanic waveguide between two vertical arrays for various waveguide parameters, along with physical explanation, will be presented.

9:00

4aSP3. Detection enhancement using multiple time-reversed guide sources in shallow water: Analysis of TREX04 data. David C. Calvo, Charles F. Gaumont, and David M. Fromm (Acoust. Div. Naval Res. Lab., Washington, DC 20375-5320)

Detection in a monostatic, broadband, active sonar system in shallow water is degraded by propagation-induced spreading. In the TREX04 experiment, performed south of the Hudson Canyon off the coast of New Jersey, 0.25 sec LFMs were transmitted with 500 Hz bandwidths chosen over a 0.5–3.5 kHz range using the NRL 64 element source-receiver array. The transmissions were then echo-repeated by a distant ship at a range varying between 1–5 km. The echo-repeating ship also transmitted one-way, 1 s, LFMs which were used as guide-source signals. These guide-source signals, which contain environmental information, form the basis of a technique for improving detection without having explicit environmental knowledge. Using an empirical-orthogonal-function representation of the set of monostatic guide-source signals, echoes were convolved with the time-reversed orthogonal functions as part of a filter bank. The result is improved probability of detection of noisy echoes using multiple guide-source signals in comparison with a baseline probability of detection using matched-filtering. In this talk, ROC curve improvement is analyzed as a function of the many parameters, e.g., target depth, source aperture, bandwidth, number of guide sources and their spatial and temporal separation. Comparison with numerical simulations is also made. [Work supported by ONR.]

9:15

4aSP4. Backpropagation image analysis of broadband decomposition of the time reversal operator (DORT) data from TREX-04. David M. Fromm, Charles F. Gaumont, Richard Menis, David C. Calvo (Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375-5320), Joseph F. Lingeitch, Geoff F. Edelmann, and Elisabeth Kim (Naval Res. Lab., Washington, DC 20375-5320)

Single-frequency DORT is a method of isolating scatterers with a multiple-source/multiple-receiver system [C. Prada *et al.*, J. Acoust. Soc. Am. **99**, 2067–2076 (1996)]. Broadband DORT overcomes problems of frequency fading and dispersion by exploiting the response of singular values across the entire signal bandwidth. In spring 2004, the Time Reversal Experiment, TREX-04, collected data for the evaluation of the broadband DORT technique. An 80 m vertical source/receiver array with 64 hydrophones was deployed in 100 m deep water in an area south of Hudson Canyon off the New Jersey coast. Sets of four, five, or six beams were transmitted at small angles from horizontal and the response from an echo repeater was recorded on 64 channels. Data was collected in 500 Hz bands between 500 and 3500 Hz with the echo repeater at ranges from 0.5 to 5.0 km and at mid-water column and near-bottom depths. Applying the

broadband DORT algorithm, the isolation of the echo repeater signal in the singular vectors and the robustness of broadband backpropagation images will be presented as a function of frequency band, signal coherence, echo repeater range and depth, noise levels, and receiver aperture. [Work supported by the ONR.]

9:30

4aSP5. Ultrasonic tomography for density gradient determination and defect analysis. Xiang Zhao and Joseph L. Rose (Dept. of Eng. Sci. and Mech., Penn State Univ., University Park, PA 16802)

Ultrasonic tomography is used to study density gradients and structural defects. In practice, it is not possible to measure an almost infinite number of projections, necessary for the transform-based techniques to produce highly accurate results. To overcome this difficulty, interpolations with respect to sample angle and projection angle based on fewer measurements are used to generate the required projection data. The number of sampled grid values necessary for displaying a well-balanced reconstructed image to reduce the aliasing distortions caused by insufficiency of the input data is then possible. For slowness reconstruction, as an example, one should constrain the time of flight at a source point to zero when using interpolation, other than using extrapolation to generate projection data to guarantee that the interpolated data around the source converges to zero. Hence no time is taken when a source and a detector overlap. Algebraic reconstruction algorithms, however, lack the accuracy when using a smaller number of measurements, and computational efficiency when using an abundant data set generated by interpolating. Results show that this approach can improve the overall accuracy and cost more effectively, thus making it a potentially, powerful tool for studying in-situ process and health monitoring. [Work supported by NSF.]

9:45

4aSP6. Implementation of a nearfield acoustic holography system based on pressure and velocity measurements. Michael C. Harris, Jonathan D. Blotter (Dept. of Mech. Eng., Brigham Young Univ., 435 CTB, Provo, UT 84602), and Scott D. Sommerfeldt (Brigham Young Univ., Provo, UT 84602)

Analytical models have indicated that a pressure and velocity-based nearfield acoustic holography (NAH) method will provide the same reconstruction resolution of current NAH methods with significantly fewer measurement locations. This would lead to a considerable savings in data acquisition time for scanning array systems and reduce the inefficiency of NAH for high spatial frequencies. The pressure and velocity-based NAH method will be introduced. Experimental results for a simply supported plate and cylinder will be presented. These results will be compared to results obtained using conventional NAH methods.

10:00–10:15 Break

10:15

4aSP7. Energetically optimal regularization in nearfield acoustical holography. Xiang Zhao (Dept. of Eng. Sci. and Mech., Penn State Univ., University Park, PA 16802)

Regularization techniques, such as singular value discarding or Tikhonov regularization, are commonly used to improve estimate of source field to be reconstructed from measured acoustic pressures at many points in nearfield acoustical holography. Theoretically, however, the regularized solution always underestimates the sound power of the real source. This paper presents an energetically optimal regularization method to solve this problem in nearfield acoustical holography by introducing a compensation factor in the generalized cross-validation function used to determine the optimal regularization parameter based on the Tikhonov regularization technique. Numerical examples show that the resulting regularization parameter will not only ensure a good fit of the predicted acoustic pressures to all measured data, but also ensure that the sound power of the equivalent source is always in agreement with the measured

sound power, thus making this regularization method more ideal and much better than others in finding a compromise between the fidelity to input data and the fidelity to source field, and yielding a robust solution to the source field. [Work supported by NSF.]

10:30

4aSP8. Multi-variate error analysis of beam-forming acoustic measurements in a wind-tunnel. Ramani Ramakrishnan, Greg Kawall, Pushpinder Bhullar (Ryerson Univ., 350 Victoria St., Toronto, ON, Canada M5B 2K3, rramakri@ryerson.ca), and Norman Ball (Inst. for Aerospace Res., NRCC, Ottawa, Canada K1A 0R6)

Source localization has been in the forefront in acoustics, since quantification of the source power and its exact location is de rigeur for successful noise control. Starting from single microphone measurements, many different methods, progressively more successful and complex, have been attempted for source localization. Recently, beam-forming methods, due to their success in underwater acoustics and architectural acoustics, have been used in wind tunnel tests. These tests have been quite extensively utilized both in automotive and aircraft tests. Beamforming techniques use an array of microphones, arranged in different patterns, to detect the source location and source power. In beamforming, the array in effect beams, not in the physical sense, towards the source to determine its position. Various parameters determine the success of the beamforming techniques. Some of these parameters are: number of microphones, microphone spacing, array pattern, source frequency and signal analysis procedures. A series of microphone array tests were conducted at the national Research Council of Canada's wind tunnel, with and without flow. The source location was known *a priori*. A multi-variate error analysis was performed to determine the importance of the different parameters. The results of the analysis will be presented in this paper.

10:45

4aSP9. Acoustical spectral analysis of a wake vortex cross-section using microphone-arrays. Hadi S. Wassaf (John A. Volpe Natl. Transportation Systems Ctr., Cambridge, MA 02142), Oliver C. Ibe (Univ. of Massachusetts, Lowell, MA 01854), and Robert P. Dougherty (OptiNav Inc., Bellevue, WA 98004)

The ability to unambiguously identify the acoustic energy generated by airplane wake vortices using phased microphone-array intensity distribution maps is hampered by the lack of knowledge of the associated non-stationary spectrum throughout the wake's duration. This ambiguity is especially pronounced when the low broadband signal to noise ratio renders the wake signal difficult to isolate from contaminating background noise. In this paper, a technique to isolate the acoustic spectral content of a cross sectional wake track (CSWT) is presented. First, image processing is used to generate a CSWT from vertical intensity distribution maps. In a second beamforming stage, the array's main-lobe traces the track to amplify the acoustic signal along the CSWT and estimate its time-varying power spectral density $P_{CSWT}(f)$ using Short Time Fourier Transform. Simultaneously, the background noise frequency-dependent upper confidence limit $CL(f)$ is estimated at each time increment by steering the array away from the wake source. The time varying $CL(f)$ is then used as an adaptive threshold function on $P_{CSWT}(f)$ in order to dynamically segment out the wake frequency bands from those of the contaminating noise sources.

11:00

4aSP10. Performance analysis of ambiguity function based broadband chirp direction of arrival estimation. Ning Ma and Joo Thiam Goh (DSO Natl. Labs., Singapore)

When estimating the direction of arrival (DOA) estimation of broadband sources with an array, the source is usually modeled as white noise. However, when the source signal has certain time-frequency structure, studies have shown that the DOA estimation performance can be improved

by using the information about its structure. In our previous work [Ning Ma and Joo Thiam Goh, "DOA Estimation for Broadband Chirp Signals," Proc. of ICASSP 2004, Vol. II, pp. 261–264], we have proposed two new methods for broadband chirp signal DOA estimation based on the ambiguity functions, namely Incoherent Broadband Chirp DOA estimation (CBD-I) and Coherent Broadband Chirp DOA estimation (BCD-C). Besides improving the accuracy of the DOA estimates, these algorithms work with signal frequencies that are higher than the array design frequency. The number of sources are also not constrained by the number of sensors. The performance of the proposed DOA estimation methods is analyzed theoretically and numerically in this work. The analysis results show that these methods can improve the output SNR and therefore the DOA estimation performance when input SNR is higher than certain level.

11:15

4aSP11. Experimental reproduction of random pressure fields in laboratory conditions. Teresa M. Bravo and Cedric Maury (Dept. de Genie Mecanique Acoustique, UTC, Ctr. de Recherche Royallieu, BP20529, 60205 Compiègne France, teresa.bravo-maria@utc.fr)

The design and the physical implementation of an experimental set-up are discussed for the off-line reproduction of random wall-pressure fluctuations with given spatial correlation characteristics. This approach could provide a cost-effective laboratory method of both reducing the variability of low frequency sound transmission measurements as well as measuring the boundary layer noise transmitted through aircraft fuselage structures. Three different types of random excitations are considered, namely an acoustic diffuse field, a turbulent boundary layer excitation and an acoustic progressive wave at grazing incidence angle. The corresponding excitations are generated in a semi-anechoic chamber using a near-field array of 4×4 loudspeakers optimally driven and located above a set of 13×16 microphones. The microphones are positioned a short distance apart from

an aluminum test panel. The optimal driving signals are determined from acoustic transfer functions measurements between the grid of microphones and the array of loudspeakers. Given this number of loudspeakers, a reasonable reproduction of the boundary-layer excitation is achieved up to about 200 Hz and a good approximation of both the acoustic diffuse field and the progressive wave is obtained over the frequency range of interest, i.e. up to about 700 Hz. [Work supported by ANVAR.]

11:30

4aSP12. Acoustic to seismic ground excitation using time reversal. Brad Libbey and Douglas J. Fenneman (Night Vision and Electron. Sensors Directorate, 10221 Burbeck Rd., Fort Belvoir, VA 22060)

Time reversal is a promising method of controlling the arrival of acoustic signals at receiver locations underwater and in biomedical tissues. This work applies the techniques in soil for landmine detection. The measurement of seismic vibrations between 50 and 300 Hz is problematic due to difficulties in transmitting energy into and through the soil from loudspeakers. The experimental setup consists of an array of loudspeakers capable of playing independent signals and a soil vibration sensor. The system response between the loudspeakers and the sensor are approximated, time reversed, and rebroadcast such that arrivals at the sensor are coincident. If the airborne and seismic paths are sufficiently complex, time reversal may help to control the energy distribution of a signal reaching the target region. However, the focusing ability is limited by an auto-correlation of the system impulse response that is inherent with time reversal processing. Simulations and experiments will be presented that demonstrate the advantages and limitations of using time reversal for controlling energy at a receiver. Results will compare vibration amplitude with and without the time reversal processing for target soil locations.

THURSDAY MORNING, 19 MAY 2005

GEORGIA B, 8:30 A.M. TO 12:00 NOON

Session 4aUW

Underwater Acoustics: Propagation: Modeling, and Experimental Results I

Kathleen E. Wage, Chair

Dept. of ECE, George Mason Univ., 4400 University Dr., Fairfax, VA 22030

Contributed Papers

8:30

4aUW1. Spectral features of sound field fluctuations in shallow water with internal solitons. Mohsen Badiey (Univ. of Delaware, College of Marine Studies, Newark, DE 19716), Valery Grigorev (Voronezh State Univ., Voronezh 394006, Russia), Boris Katsnelson (Voronezh State Univ., Voronezh 394006, Russia), and James Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Mode coupling of acoustic wave propagation in shallow water, in the presence of internal soliton is considered. It is shown that quasiperiodical behavior of temporal variability of the received signal (arising from the motion of soliton packets) provides some peaks in frequency spectrum. Positions of these peaks in spectrum are determined both by the space scales of interference beating (ray cycles) and velocity of solitons. Results of theoretical and numerical modeling as well as related experimental setup is discussed. [Work supported by ONR, RFBR and CRDF.]

8:45

4aUW2. Influence of random nonlinear internal wave parameters on resonant acoustic mode coupling. Scott D. Frank (Dept. of Mathematics, Marist College, Poughkeepsie, NY 12601) and William L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180)

Shallow water transmissions sometimes suffer anomalous amplification or loss caused by mode coupling due to the presence of nonlinear internal wave packets. The possibility of coupling between two acoustic modes is specified by an internal wave-acoustic resonance condition, which relates acoustic wavenumber differences to peak locations of the packet spectrum. This mechanism is critical in a recent analysis [S. D. Frank *et al.*, J. Acoust. Soc. Am. **116**, 3404–3422 (2004)] of some broadband intensity variations from the SWARM 95 experiment. That investigation used regularly spaced, evenly-sized nonlinear internal wave packets and it is important to determine how random variations in the wave widths and inter-wave spacings affect propagation results. An analytic expression

for the wavenumber spectrum of an idealized packet is obtained in terms of these parameters. This expression predicts spectral peak locations for parameters from the SWARM 95 model that correspond with observed data. Estimates of peak location statistics arising from random packet parameter variations are analyzed computationally and analytically. Variability of internal wave spectral peak locations is correlated with broadband intensity fluctuations obtained from parabolic equation computations. [Work partially supported by ONR.]

9:00

4aUW3. Influence of internal wave fluctuations on acoustic underwater propagation. Tatiana A. Andreeva, William W. Durgin, and Stefanie E. Wojcik (Mech. Eng. Dept., Worcester Polytechnic Inst., 100 Institute Rd., Worcester, MA 01609)

The work presents a ray acoustic based analysis of the effect of ocean internal waves and its fluctuations on acoustic wave propagation. In the present work, it is shown that consideration of only perfect internal waves simplifies the real problem, disregarding the fluctuations of internal waves. The focus of the paper is to study numerically the influence of internal wave fluctuations and the role of an initial ray angle onto underwater acoustic propagation. In the present formulation the eikonal equation is considered in the form of a second order, nonlinear ordinary differential equation with harmonic excitation due to internal wave. The harmonic excitation is taken imperfect, i.e., with a random phase modulation due to Gaussian white noise. The amplitude and wavelength of the acoustic waves are used as the principle signal characteristics in bifurcation analysis. The regions of instability are identified using the bifurcation and phase diagrams. The effect of internal waves phase modulation is demonstrated by means of comparison of the numerical data obtained in the present work with data, obtained perfect internal wave. The preliminary analysis shows very strong dependence of acoustic propagation on intensity of fluctuations.

9:15

4aUW4. Influence of internal waves on vertical coherence of sound propagation in the East China Sea. Jie Yang and Peter H. Rogers (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405)

Literature of the last two decades has shown that internal waves are a major factor determining sound field fluctuations [Zhou *et al.*, J. Acoust. Soc. Am. **90**, 2042–2054 (1991); Lynch *et al.*, *ibid.* **99**, 803–821 (1996)]. Due to its time-dependent and anisotropic nature, internal waves can cause both temporal and spatial variations in the sound speed profile and hence, the acoustic field. In this paper, a model is built to study the range and azimuthal dependence of wave propagation in the presence of internal wave activity. The results are based on an *in situ* experiment ASIAEX in the East China Sea where intense tidal driven internal waves were recorded. The model explains the observed decorrelation after long range propagation (20 km) and in certain directions relative to the internal wave wavefront.

9:30

4aUW5. Intensity and time-angle fluctuations of high frequency signal propagation in shallow water. Mohsen Badiy, Jing Luo, and Aijun Song (Univ. of Delaware, College of Marine Studies, Newark, DE 19716)

Several high frequency experiments in recent years have been directed towards assessing signal variability as a function of channel environmental parameters. In this paper, results from one of these experiments in a shallow water region are shown. Acoustic communication coded signals were transmitted between a source and two receiver arrays separated 1 and 2 km respectively while detailed measurement of the channel environmental parameters were recorded. Channel impulse response function and the correlation time of the channel are obtained. Ray theory is used to analyze and interpret the experimental data. Arrival time-angle fluctuations were found to be directly correlated with the environmental variability due to ocean dynamics in the experiment region.

4aUW6. Mean and variance of the 3-D field forward propagated through random internal waves in continuously stratified deep and shallow water ocean waveguides. Tianrun Chen, Purnima Ratilal, and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

A general modal solution for the mean and covariance of an acoustic field propagating through 3-D random inhomogeneities in an ocean waveguide [Ratilal and Makris, J. Acoust. Soc. Am. **114**, 2428 (2003)] is applied to propagation through a general 3-D internal wave field. Our previous two-layer internal wave parameterization is extended to continuous stratification and applied in both continental shelf and deep ocean waveguides. Physical and statistical properties of the internal waves are modeled with the Garret-Munk and other theoretical and empirical formulations. The mean and variance of the forward field propagating through internal waves is expressed in terms of the stochastic moments of the medium's scatter function density. This is formulated with the Rayleigh-Born approximation to Green's theorem in terms of the medium's compressibility and density fluctuations. In this way, the effect of both internal-wave sound speed and density variations on the forward propagated field are quantified. Density variations are shown to greatly affect high order modal propagation in some cases. Two-dimensional models of propagation through random internal wave fields are shown to become inaccurate after exceeding threshold ranges where 3-D scattering becomes important.

10:00

4aUW7. A mean field transport theory for long-range ocean acoustics. Michael Wolfson and Frank Henyey (Univ. of Washington, APL, 1013 NE 40th St, Seattle, WA 98105-6698)

Voronovich has found from an ocean model that the mean acoustic wavefield at low frequency persists after multi-megameter propagation in the deep ocean. The common belief has been that only the incoherent field survives long propagation distances. His results, if true in the real ocean, have important consequences to extracting signals from long-range propagation. The results of an alternative ocean model, for which the mean field is evaluated using transport theory, are presented, and compared to his results. Our model consists of a single background sound speed field superposed with sound speed fluctuations due to internal waves. We present results for sources used in recent experiments such as LOAPEX (75 Hz with about 30 Hz bandwidth).

10:15–10:30 Break

10:30

4aUW8. Calculation of amplitudes of acoustic normal modes from the reciprocity principle. Oleg A. Godin (CIRES, Univ. of Colorado and NOAA/Environ. Technol. Lab., Mail Code R/ET-0, 325 Broadway, Boulder, CO 80305)

Recently, J. D. Achenbach [J. Acoust. Soc. Am. **116**, 1481–1487 (2004)] put forward, on heuristic grounds, an elegant technique to calculate amplitudes of guided waves generated by mechanical loading in a range-independent elastic medium. The technique is based on application of the reciprocity principle and allows one to solve the problem without use of integral transforms. In this paper, the technique is applied to acoustic waveguides in a layered fluid. By taking into account continuous spectrum of the field, which was disregarded by Achenbach, a mathematical justification of the technique is obtained. The technique is shown to be exact. It is further extended to enable calculation of excitation coefficients of modes of both discrete and continuous spectra by a given sound source. The results are shown to be identical to those derived with the traditional, more cumbersome methods that rely on integral transforms. [Work supported by ONR.]

10:45

4aUW9. Moving ship tomography a first look at BASSEX 04. Kevin D. Heaney (OASIS Inc., Falls Church, VA 22044) and Arthur B. Baggeroer (MIT)

During the Basin Acoustic Seamount Scattering Experiment (BASSEX04) a line array was towed between two source moorings in the North Pacific separated by 500 km. The moorings were part of the SPICE04 test led by Scripps Institution of Oceanography. Each mooring had two broadband sources (center frequency 250 Hz) at depths of nominally 800 m and 3000 m. Thirty transmissions were recorded as the ship passed south along the line bisecting the two source moorings. The sources transmitted simultaneously (for each source). In order to break the conical ambiguity of the arrivals from each source, the array was turned 20 degrees for each reception. One set of sources used coded m-sequences with different laws, permitting separation of the signals by signal processing. Comparisons of the received signals will be compared with initial propagation model results using the measured sound speed field from the ship. An approach to $N \times 2-D$ tomography (where $N=30$) will be investigated. [Work supported by ONR Ocean Acoustics.]

11:00

4aUW10. Basin Acoustic Seamount Scattering Experiment (BASSEX 04) II. Downslope propagation. Kevin D. Heaney (OASIS Inc., Falls Church, VA 22044), Arthur B. Baggeroer (MIT), Kyle M. Becker (ARL-Penn State Univ.), Eddie Scheer, and Keith Vonderheydt (Woods Hole Oceanogr. Inst.)

During the Basin Acoustic Seamount Scattering Experiment (BASSEX04), the Five-Octave Research Array (FORA) was deployed near the North Pacific Acoustics Laboratory (NPAL) bottom mounted tomography source just off the coast of Kauai. Receptions from the Kauai source were taken at ranges from 2 km out to 500 km. The data was taken to examine the affects of seafloor interaction on downslope propagation (propagation down both steep and shallow slopes), shallow water range-dependent 3-D propagation and conversion from shallow water propagation to deep water propagation. Conventional and adaptive beamforming results will be presented as well as comparisons between measured responses and PE modeled responses. Geo-acoustic inversions will be performed with the broadband source in end-fire and broadside orientations of the array. [Work supported by ONR Ocean Acoustics.]

11:15

4aUW11. Forward scattering from the Kermit-Roosevelt Seamount complex during the SPICEX-LOAPEX-BASSEX experiments. Arthur B. Baggeroer, Joseph Sikora III (Massachusetts Inst. of Technol., Rm. 5-206, Cambridge, MA 02139), Kevin Heaney (Oasis Corp., Lexington, MA 02173), Edward K. Scheer, Keith von der Heydt (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Kyle Becker (Pennsylvania State Univ., State College, PA)

The SPICEX-LOAPEX-BASSEX experiments were executed in the Northeast Pacific to examine several long range, low frequency propagation phenomena. Low frequency sources centered at 68 and 75 Hz with nominal bandwidth of 30 Hz and a source at 250 Hz with 100 Hz bandwidth were deployed transmitting orthogonal M sequences and FM chirps. A 64 element towed array cut for 250 Hz from Penn State Univ. was the

primary multichannel receiving system. This presentation concerns the BASSEX study of low frequency scattering around the Kermit-Roosevelt Seamount complex which shoals to 900 m near 39 N and 145 W. We examine (i) the forward scattering shadow, (ii) any patterns of horizontal refraction as a function of source and receiver range from the seamounts, (iii) any backscattering from the seamounts, and (iv) the modal content of the signals by travel time methods. In addition, directional spectra of ambient noise were measured.

11:30

4aUW12. Three different directions of sound propagation in the China South Sea. Ma Li (Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, China)

In June 2004 three directions of sound propagation experiments were done in the China South Sea. One direction is equally sea depth. The second direction is slope. The third direction is largest gradient in sea depth direction. In the mean time three temperature chains have been applied for measuring the internal wave. During each temperature chain their distances are about 1000 m. The experiment results show that the packet of solitary waves has been found and each direction of sound propagation has different transmission loss. In the equally sea depth direction the sound propagation has the smallest transmission loss. And in the largest sea depth direction the sound propagation has the largest transmission loss. By three temperature chains data the packet of solitary waves propagation velocity and direction may be measured. And through the packet of solitary waves propagation direction the packet of solitary waves may be induced by the West Sand Island. In the end the couple normal mode program has been used to simulate the sound propagation which is affected by the packet of solitary waves. The simulating results are compared to experimental data.

11:45

4aUW13. Long-range acoustic transmission in a deterministic range-dependent ocean: Contribution from interference of near-axial waves. Natalie S. Grigorieva and Gregory M. Fridman (Dept. of Appl. Mathem. and Mathem. Modeling, St. Petersburg State Marine Tech. Univ., 3 Lotsmanskaya Str., St. Petersburg, 190008, Russia, nsgrig@natalie.spb.su)

The propagation of energy along the waveguide axis cannot be described in terms of geometrical acoustics because of the presence of cusped caustics repeatedly along the axis. In neighborhoods of these cusped caustics a very complicated interference pattern is observed. Neighborhoods of interference grow with range and at long ranges they overlap. For an arbitrary range-independent ocean, it was shown in [N. S. Grigorieva and G. M. Fridman, *J. Comp. Acoust.* **12**, 127-147 (2004)] that the interference of wave fields corresponding to near-axial rays results in a diffractive (as opposed to geometrical acoustics) component of the field which can be viewed as distinct, "axial" wave and effectively calculated using its integral representation. In this paper the integral representations of the axial wave in the frequency and time domains are obtained for a range-dependent ocean. In the frequency domain this representation has the form of a linear superposition of the solutions of the Helmholtz equation that are concentrated in a neighborhood of the sound-channel axis. The weight function is selected in such a way that the localization principle holds. The axial wave is simulated for a deterministic model of a range-dependent ocean corresponding to AET experiment. [Work supported by ONR Global.]

4a THU. AM

Session 4pAA**Architectural Acoustics and Noise: Soundscapes from an Architectural Viewpoint**

Brigitte Schulte-Fortkamp, Chair

*Technical Univ. Berlin, Inst. of Technical Acoustics, Secr TA 7, Einsteinufer 25, Berlin 10587, Germany***Chair's Introduction—2:15*****Invited Papers*****2:20****4pAA1. The soundscape as a tool for urban planners and town designers.** Catherine Semidor (GRECO, EAPBx, Domaine de Raba F-33400 Talence, France, catherine.semidor@bordeaux.archi.fr)

Urban planners and other town designers need information to improve the environmental quality of cities, for instance the components of the urban soundscape. To use the latter as a decision-making tool in a planning process it is necessary first to link objective and subjective acoustic criteria with architectural parameters and then to transmit the observed characteristic relations in a way which designers will understand. One of these ways is to make use of spatial references (streets' shape, facades' material, etc.) to explain the impact of the urban morphology, for example, on the sound environment. That is why, for many years, surveys were conducted in a lot of different areas. This paper deals with some comparative results obtained in different European towns.

2:45**4pAA2. Phase characterization of soundscapes.** Vincent Gibiat, Abril Padilla (PHASE, Univ. Paul Sabatier, 118 route de Narbonne, 31062 Toulouse, France, gibiat@cict.fr), Valentin Emiya (ENST, 75014, Paris, France), and Lionel Cros (ENSIETA, 29806 Brest Cedex 09, France)

Signal processing of soundscapes remains far from the auditive perception. It is particularly true when sounds recorded inside railway stations, are compared with the most accurate spectrographic representations. Such architectural realizations provide a soundscape between closed and opened spaces. As microphones used to discriminate sound sources as well as low level sounds use diffractions properties of radiation, as time discrimination between near sounds is efficient through time domain perception, as filtering often introduces phase distortions, it has been decided to explore a new kind of signal representation based on the phase of the Short Time Fourier Transform. After a basic presentation of the numerical tool named SAFIR for Spectrography in Amplitude and Frequency, Instantaneous and Reallocated, examples will be given that first show the differences between closed and opened spaces, then the differences between the soundscape inside a train and inside a station. As this work has been realized with sounds recorded for a later musical realization that deals to reconstruct both a new realistic or imaginary coherent soundscape, it has been essential to extract the pertinent, coherent, details as magnification of time or space opposed with the incoherent urban noise. Finally SAFIR will be compared with the composer work.

3:10**4pAA3. How does hearing get into architectural design?** Juergen Bauer (Architect, Mierendorffstr. 11, 10589 Berlin, Germany, jbauer.berlin@tiscali.de) and Brigitte Schulte-Fortkamp (Tech. Univ. Berlin, D-10587 Berlin, Germany)

Serious soundscape research needs an interdisciplinary approach to architecture. The design and the perception of architecture are dominated by the visual senses. Architects developing a complex building may be advised by a specified acoustic engineer. However the support provided by the acoustic engineer focuses less on creating acoustic qualities than on minimizing acoustic problems and noise effects. How about the sense of hearing of the architect? Designing architects have learned to shape 3-D. They have a precise vision of space and proportion. They know about light and shadow, about fabric and color. This profound world of vision is the basis of their inspiration, and they would be grateful to get more consciously in touch with the world of hearing not in terms of creating extraordinary sound-effects but in terms of imaginative listening. The visual phenomena architects are dealing with offer an approach to the world of hearing or rather analogies from watching towards listening. Light, shadow, fabric and color are no more matters only of sight but subject to audiovisual awareness. What acoustic processes might be most relevant to designing architects? What acoustic events and impacts should be specified in the education of design?

3:35**4pAA4. Soundscape perception of an urban park under flight paths of Naples airport.** Giovanni Brambilla (CNR-Inst. of Acoust., via del Fosso del Cavaliere 100, 00133 Rome, Italy, giovanni.brambilla@idac.rm.cnr.it), Leda De Gregorio, and Luigi Maffei (Second Univ. of Naples, 81031 Aversa, Italy)

The Capodimonte historical urban park is the largest green area in Naples, located in the North and under the flight-paths of the city airport. Noise measurements, taken in the morning on two weekdays and one week-end in different sites inside the park, have shown LAeq levels, excluding aircraft noise, lower than 50 dB(A) and up to 55 dB(A) on Sunday morning near the museum. During the noise measurements one aircraft flyover occurred approximately every five minutes, producing an increase of LAeq in the range

5-15 dB(A) and LAmax levels up to 90 dB(A). Interviews have been carried out on a sample of subjects present in the park while the noise measurements were taken. By means of a questionnaire, they were asked to rate different factors describing the environment and influencing their experience to enjoy the park. The analysis of the subjective data collected is in progress and the preliminary outcomes show that the reported annoyance due to aircraft noise is much lower than expected by the noise levels, most likely due to synergetic effects of the other environmental factors.

4:00

4pAA5. Community design with soundscape in mind. Bennett M. Brooks (Brooks Acoust. Corp., 27 Hartford Turnpike, Vernon, CT 06066, bbrooks@brooksaoustics.com)

Community standards for sound are usually defined by limits on noise emitters. These limits may be absolute in level, or relative to a background level. Most often a limit is given as an overall A-weighted level. Occasionally, limits are placed on octave band spectrum levels or on tonal content. The time function of noise is generally addressed only broadly, defined by terms such as impulse, intermittent, or continuous. The character of sound in the community is addressed in policy or law by vague definitions which prohibit sounds which are obnoxious or that cause a nuisance. The positive aspects of sound in communities are usually left to chance. Recently, one summer resort community initiated a policy which promotes outdoor entertainment in one district, while requiring quiet in another. This unique policy balances lifestyle and economic interests, and was developed through an inclusive community-based public process. The question of how the community-based policy development process may be used to enhance the soundscape of a targeted area is explored. The implications of such a policy process for the designers of buildings and public spaces are discussed.

Contributed Papers

4:25

4pAA6. Aesthetic transformations of soundscapes as a basis for architectural design. Gary Siebein (Univ. of Florida School of Architecture, P.O. Box 115702, Gainesville, FL 32607)

A graduate design studio in acoustics examines sound and silence as manifest in urban and rural exterior soundscapes as well as interior soundscapes within performance venues as potential theoretical and formal bases for architectural form. The soundscape analyses were composed of a series of quantitative and qualitative studies of real and virtual environments such as the actual project site in downtown Orlando and the interior soundscape of the Stats Opera in Vienna in the 19th century. Transformations of the soundscape analysis served as generative ideas for the basis of architectural interventions in complex sites and programs for a multi-venue performing arts center.

4:40

4pAA7. Bow bells—creating a sound community. Brenda Kiser and David Lubman (David Lubman & Assoc., 14301 Middletown Ln., Westminster, CA 92683)

For centuries, the sound of ringing church bells has given communities a sense of identity. Church bells have long been used to announce significant events to the community, such as births and deaths, to warn of impending danger, to call the faithful to church and to announce the curfew. This paper focuses on one prominent example of a sound community—the church of St. Mary-le-Bow in the City of London, the location of the famous Bow bells. These bells have featured in London's folklore and history since the 14th century and, to this day, identify those born within the sound of Bow bells as Cockney. This sound neighborhood was extended far beyond its natural borders when a recording of the bells was played by the BBC's World Service during World War II in broadcasts to occupied Europe. Sound examples are given. A short history of European church bells is presented. A cognitive scale is proposed, ranging from belonging to alienated, delineating psychological reactions of those within a sound neighborhood.

THURSDAY AFTERNOON, 19 MAY 2005

PLAZA C, 1:25 TO 4:55 P.M.

Session 4pAB

Animal Bioacoustics: Tools for Animal Bioacoustics: New Designs and Directions II

Susanna Blackwell, Chair
120 Tamarack Dr., Aptos, CA 95003

Chair's Introduction—1:25

Contributed Papers

1:30

4pAB1. Autonomous remote humpback whale acoustic monitoring devices. Whitlow Au, Alison Stimpert, Marc Lammers (Marine Mammal Res. Program, Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734), and Niklas Johnson (Lund Univ., Lund, Sweden)

Four acoustic monitoring devices controlled by a CF-1 Persistor Instrument Inc. microcontroller board were developed to monitor the chorusing sounds of humpback whales in their wintering grounds off the

island of Maui. The units were programmed to turn on at the same time and sample chorusing sounds at a sampling rate of 3.33 kHz for 3 minutes every hour. The CF-1 was chosen because of its low power requirement of 40 ma while collecting data during the 3-min collection period, 3 ma while sleeping and 96 ma while storing the data onto a 2 Gbyte compact flash card. At this sampling rate, sampling period duration, and interval between sampling period, each unit has sufficient power to last 30 days when powered from a 12 Ahr lead acid battery. The 2 GB memory allows for 34 days of data collection. Four of these units were deployed along the west

shore of Maui at locations separated by 15 to 20 miles. The amplitude of chorusing by singing humpback whales and the temporal characteristics at the different locations were obtained with these sensors during the 2005 humpback whale season.

1:45

4pAB2. Monitoring ecologically important fish sounds: The characteristics and diel trends of sounds produced by three species of Pomacentrids. T. Aran Mooney, Marc O. Lammers (Hawaii Inst. of Marine Biol., Univ. of Hawaii, Kaneohe, HI 96744), Pedro A. Santos (Univ. of the Azores), and Paul E. Nachtigall (Univ. of Hawaii)

Many fish species make sounds. These sounds can be recorded, characterized, identified to species, correlated with behavior, and then monitored for temporal and behavioral trends. We developed a two-part approach to accomplish these tasks. The first is a Portable Underwater Acoustic (PUA) device, composed of a two-hydrophone, handheld, digital recorder. This system is used to identify fish sounds that can then be correlated with specific behaviors. The second is a single-hydrophone, autonomous recording system that is placed on the seafloor to record the ambient sound field at set temporal intervals. This presentation describes both systems and how they are being applied in the field to study three species of fish: *Abudefduf luridus*, in the Azores, Portugal, and *Dascyllus albisella*, and *Abudefduf sordidus*, in Hawaii, USA. *A. luridus* sounds were usually produced in a doublet, 100 ms in duration, a peak frequency of 400–450 Hz, and associated with aggressive displays. Sounds produced by *D. albisella* were typically 2–4 s, with a peak frequency of 300–400 Hz and a bandwidth of up to 6 kHz. *A. sordidus* produced 100 ms broadband clicks, up to 5 kHz bandwidth, in aggressive chases. After identification, sounds were associated with behaviors and ecologically important events, and monitored for diel trends in their production.

2:00

4pAB3. Application of a diver-operated 4-channel acoustic/video recording device to study wild dolphin echolocation and communication. Michiel Schotten, Marc O. Lammers (Marine Mammal Res. Program, Hawaii Inst. of Marine Biol., 46-007 Lilipuna Rd., Kaneohe, HI 96744, micschotten@hotmail.com), Ken Sexton (The Sexton Co. LLC, Salem, OR 97301), and Whitlow W. L. Au (Hawaii Inst. of Marine Biol., Kaneohe, HI 96744)

Despite decades of research on captive dolphins, there still is much to learn about dolphin echolocation and communication in the wild. This is mainly due to the difficulties associated with recording dolphin signals accurately in the wild, such as their broadband (with frequencies over 200 kHz) and directional nature, the difficulty of identifying vocalizing dolphins, and being confined to equipment aboard a vessel. To resolve these difficulties, an underwater portable battery-operated digital recording device was developed that is capable of simultaneously capturing acoustic signals up to 220 kHz on 4 channels as well as video. It consists of a custom-made underwater housing, containing two batteries, a digital camcorder, and an acoustic recording unit with signal conditioning board, 4-channel data acquisition card, PC/104-plus single board computer and power supply, and a notebook hard disk. Attached to the housing is a 4-hydrophone star array. Post-recording, acoustic signals recorded within 15 m are accurately localized in 3-D and attributed to individual dolphins on the video. Thus, acoustic signal features are correlated with different echolocation behaviors and communicative signal exchanges among dolphins are assessed. The system has been successfully field tested with spinner dolphins (*Stenella longirostris*) along the leeward coast of Oahu, Hawaii.

2:15

4pAB4. A new acoustic bat detector. Matthew Heavner (Univ. of Alaska, Southeast, 11120 Glacier Hwy., Juneau, AK 99801, matt.heavner@uas.alaska.edu)

A new small, low-power, long-duration, field-deployable computer for the acoustic monitoring of bats in Southeast Alaska is being developed. The most recent study of bats in the region, reported by Parker *et al.*, (1997), found the geographic range of four bat species to have northern distribution limits in Southeast Alaska. (A fifth species, *E. fuscus*, goes much farther north.) The relationship of the bats to forest management practices, the habitat usage of the bats, and population size and trends are all very poorly known for bats in Southeast Alaska. Long duration monitoring of several different types of area (such as old-growth versus recently logged forest) will provide knowledge to improve management practices in regards to bat ecology in Southeast Alaska. With the motivations just described, the hardware, development methods, and analysis software designed to develop an improved detector are presented. [Work supported by Alaska Department of Fish & Game.]

2:30

4pAB5. Automatic detection of microchiroptera echolocation calls from field recordings using machine learning algorithms. Mark D. Skowronski and John G. Harris (Computational Neuro-Eng. Lab, Univ. of Florida, Gainesville, FL 32611, markskow@cnel.ufl.edu)

The authors have recently presented experimental results of applying machine learning algorithms, used extensively in human automatic speech recognition research (ASR), to automatic species identification of echolocating bats [Skowronski and Harris, *J. Acoust. Soc. Am.* **116**, 2639 (2004)]. The results of those experiments demonstrated that *frame-based* classification, preferred in ASR, out-performs *holistic* classification typically employed in automatic echolocating species identification. The authors have extended the paradigm of machine learning algorithms to the related problem of bat call detection. A robust automatic bat call detection algorithm, to replace hand labeling, is required for two reasons: (1) for real-time species identification in the field, and (2) because hand labeling is subjective, tedious, slow, and error-prone. The current experiments compare various frame-based features (log energy, pitch estimates, pitch slopes) with several models of detection (matched filters, Gaussian mixtures, decision trees). Detector sensitivity and specificity are quantified for comparison using hand-labeled calls, with considerations of classification requirements for detected calls. That is, a detector is not penalized for including a short segment of background signal before and after a hand-labeled call. The results demonstrate the superior performance of the frame-based features and machine learning detection algorithms compared to conventional features and detection algorithms.

2:45

4pAB6. Detection of FM signals in the presence of non-Gaussian noise. Ildar Urazghildiiev, Kathryn Cortopassi, and Christopher Clark (Bioacoustics Res. Program, Cornell Lab. of Ornithology, 159 Sapsucker Woods Rd., Ithaca, NY 14850)

In bioacoustics, the problem of detecting frequency-modulated signals in the presence of non-Gaussian noise is of great interest. Matched filters (MF) are often chosen over frequency-specific energy (FSE) detectors because of their improvement in both signal gain and target specificity. Under non-Gaussian conditions however, MF detectors do not ensure acceptable trade-off between false alarm and missed detection rate. To decrease false alarm rate, we propose a two-stage detection technique. First, a MF is applied, with threshold prescribed by the acceptable false alarm rate, to generate candidate detections. Second, a signal recognition (SR) algorithm is applied to the candidates. The SR algorithm estimates modulation parameters from the signal spectrogram, and a detection decision is made based on how well parameters match a criterion set for the signal of interest. We applied this technique to right whale contact calls recorded in Cape Cod Bay. Results demonstrate that the MF-SR technique decreases false alarm rate by 3–5 times, while not increasing missed detection rate

from that of a MF only. These results were compared to those of a FSE detector followed by the same signal recognition and decision stage (FSE-SR), to determine if comparable performance could be achieved without the MF first stage.

3:00–3:10 Break

3:10

4pAB7. A computational model of echolocation: Restoration of an acoustic image from a single-emission echo. Ikuo Matsuo and Masafumi Yano (Res. Inst. of Elec. Commun., Tohoku Univ., Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan, matsuo@riec.tohoku.ac.jp)

Bats can form a fine acoustic image of an object using frequency-modulated echolocation sound. The acoustic image is an impulse response, known as a reflected-intensity distribution, which is composed of amplitude and phase spectra over a range of frequencies. However, bats detect only the amplitude spectrum due to the low-time resolution of their peripheral auditory system, and the frequency range of emission is restricted. The amplitude spectrum varies with the changes in the configuration of the reflected-intensity distribution, while the phase spectrum varies with the changes in its configuration and location. Here, by introducing some reasonable constraints, we propose a method for restoring an acoustic image from the echo produced by a single emission. The configuration is extrapolated from the amplitude spectrum of the restricted frequency range by using the continuity condition of the amplitude spectrum at the minimum frequency of the emission and the minimum phase condition. The determination of the location requires extracting the temporal changes of amplitude spectra. For this purpose, the Gaussian chirplets with a carrier frequency compatible with bat emission sweep rates was used. The location is estimated from the temporal changes of the amplitude spectra. This method can determine the acoustic images of objects.

3:25

4pAB8. Broken line 3-dimensional sperm whale diving trajectory reconstruction using passive acoustics on a single hydrophone. Christophe Laplanche, Olivier Adam, Maciej Lopatka, and Jean-François Motsch (Université Paris XII, 61 av. du Gal de Gaulle, 94010 Créteil, France)

Sperm whales make deep dives to hunt. A dive, lasting 45 minutes on average, is composed of a vertical descent to the prey layer depth, the properly so called hunt (at a quasi constant depth) inside the prey layer, and a vertical ascent back to the sea surface. Sperm whales make series of echolocation signals (*clicks*) during the two first stages. The sea surface/bottom click echo detection and delay measurements then make possible the sperm whale range/depth estimation during these stages, by passively using a single hydrophone. The vertical, rectilinear sperm whale trajectory during the first stage is unambiguously estimated from the echo delays. The sperm whale trajectory can also be reconstructed during the second stage, from the sperm whale range variations only, even when not detecting sea bottom click echoes. These range variations strongly suggest the sperm whale trajectory to be a broken line (e.g., a 2-piece line, 600 m straight ahead, 85 degree bend, 1000 m straight ahead). Assuming a vertical silent reascent, the rough hydrophone-relative 3-dimensional sperm whale trajectory can then be reconstructed for the complete dive, by using a single hydrophone.

3:40

4pAB9. Transmitting beam patterns of an echolocating bottlenose dolphin. Thomas G. Muir, Tobias J. Lemerande, Steven R. Baker (Phys. Dept., U.S. Naval Postgrad. School, Monterey, CA 93943, tmuir@olemiss.edu), and Samuel H. Ridgway (U.S. Navy Marine Mammal Program, SPAWARESCEN San Diego, CA 92152)

Measurements on a free-swimming subject, echolocating under a simple go paradigm, were conducted at SSC San Diego, with a linear array of seven simultaneously and individually processed wideband hydrophones (response to 400 kHz), arranged either vertically or horizontally, and cen-

tered on a small underwater video camera, with the video output synchronized with the recording of acoustic data. The measurement apparatus itself served as the test target. Lowering it into the water provided the cue for the blindfolded subject to locate its position in the test pen, swim to the apparatus, and touch it with the tip of its rostrum; whereupon the trainer provided a bridge signal, indicating reward due. During this process, acoustic beampattern measurements were made on echolocation clicks, as a function of frequency, which support previous, non-simultaneous, statistical beampattern measurements. Early in the experiment, the subject emitted echolocation clicks that peaked at 78 kHz, with -3 dB beamwidths of 8 to 10 degrees; but also containing high frequency components (above 135 kHz), with beamwidths that narrowed with increasing frequency, ranging to only a few degrees around 300 kHz, far beyond the subjects hearing range. After several days, the task appeared to become easier and there was much less high frequency content, but the click repetition frequency increased. Wideband noise was then introduced into the test pen, at frequencies below 135 kHz; and as the task became more difficult, the subject resumed transmitting clicks with high frequency components, although the peak frequency remained at 78 kHz, and click rates again increased, to much higher values.

3:55

4pAB10. A juvenile salmon acoustic tracking system (JSATS) for the Columbia River estuary. Lynn McComas (Natl. Marine Fisheries Service, 3305 E. Commerce St., Pasco, WA 99301) and Alex Easton (SAIC, Lynnwood, WA 98037)

The Juvenile Salmon Acoustic Tracking System (JSATS) was developed as a means to estimate the downriver survivability of sub-yearling salmon in the Columbia River. The system consists of a bottom-mounted sonar array which detects signals from acoustic projectors that are surgically implanted in juvenile salmon. JSATS is the result of a development program which began in 2001 when Northwest Fisheries Science Center National Marine Fisheries Service (NMFS) and Science Applications International Corp. (SAIC) developed the top level, biology driven, requirements for the system. SAIC then conducted environmental surveys to characterize the physical and acoustic environments in which JSATS would operate. Battelle PNNL developed the acoustic projector used in JSATS. The signaling characteristics of this projector are driven by the requirement that it not affect the behavior of the 90 mm salmon in which it is implanted, and that it operate continuously for 30 days. SAIC developed the detection array portion of JSATS. The design was driven by the requirements that it be bottom mounted, cabled to shore, and successfully withstand high current and moving sand waves. NMFS and SAIC jointly developed the shore stations and the installation/recovery procedures. Installation of an array spanning the Estuary is planned for April 2005.

4:10

4pAB11. Testing acoustic fish deterrents for use under-ice in arctic lakes. Roberto G. Racca (JASCO Res. Ltd, 2101-4464 Markham St., Victoria, BC, V8Z 7X8 Canada, rob@jasco.com)

Acoustic deterrent technologies can be used in aquatic settings in lieu of physical barriers to keep fish away from potentially harmful industrial operations. The goal of this study was to determine the efficacy of portable, temporary acoustic deterrents as a means of excluding fish from the neighborhood of sub-bottom detonation activities associated with seismic exploration under ice in arctic lakes. In October 2003 trials were conducted in Dolomite Lake near Inuvik, Northwest Territories, Canada on indigenous fish species. Groups of fish were equipped with orally inserted ultrasonic tags, placed in a large experimental net pen and monitored using an acoustic tracking system that produced a detailed three-dimensional swimming pattern for each subject, thereby revealing any behavioral responses. A flex-tensional broadband sound projector driven by digitally synthesized signals was tested as deterrent, and real-time monitoring at two hydrophone sites was used to estimate the local level of insonification throughout the pen volume. Although the study did not identify an overall effective deterrent, sufficient indications of response were observed to

support the future testing of a louder projector capable of emitting a tonally modulated sound pattern at frequencies from about 100 Hz to a few kHz. [Work supported by ESRF (esrfunds.org).]

4:25

4pAB12. Design of an oscillating flow test chamber for modeling the fish ear. Charlotte Kotas, Peter Rogers, and Minami Yoda (Mech. Eng. Dept., Georgia Inst. of Technol., Atlanta, GA 30332-0405, gtg227d@mail.gatech.edu)

The acoustically induced oscillating flow fields inside the fish ear, specifically near the otoliths, may provide clues about how fish directionalize underwater sound sources. A new experimental test apparatus has been constructed to study oscillatory and steady streaming flows around model otoliths. The flows are driven by a shaker-piston assembly attached to a 250 mm diameter, 450 mm tall cylindrical test chamber filled with liquid to a depth of 400 mm. This apparatus is capable of generating flow oscillating at frequencies of 1 to 30 Hz with displacements up to a few millimeters within the test chamber. Theoretical predictions of the fluid-borne vibrations in a fluid-filled tube in a vacuum suggest that the flow will have a uniform sinusoidally oscillating velocity field over the central portion of the test section. The flow is visualized in the optically accessible test chamber using small tracer particles; flow velocity fields are measured using particle image velocimetry (PIV). The apparatus is compatible with various working fluids of different viscosities such as water, glycerin and

silicon oils to extend the range of attainable Reynolds numbers. Initial results for oscillating flows around basic "otolith" geometries such as spheres and spheroids will be shown. [Work supported by ONR.]

4:40

4pAB13. Biomass assessment in commercial catfish ponds. James Chambers (Natl. Ctr. for Physical Acoust., The Univ. of Mississippi, University, MS 38677), C. Douglas Minchew (Mississippi State Univ., Stoneville, MS 38776), and Rachel Beecham (Mississippi Valley State Univ., Itta Bena, MS 38941)

With increasing seafood demand aquaculture is poised to become a major growth industry in the United States in the 21st century. In particular channel catfish represent an approximately \$260 million industry in Mississippi with strong growth potential. A major portion of the costs associated with raising channel catfish are related to the cost of feed and aeration which are directly related to the total number of fish being raised in each pond. Crop insurance and bank loans are also contingent upon accurate population estimates. A high frequency, horizontally scanning, active pulse echo sonar system, the Aquascanner, has been developed to estimate pond populations. Commercial catfish ponds are typically 2 to 10 acres and, unlike many offshore fisheries, have fairly shallow depths of 1 to 2 m. The system components and its use will be presented along with results from field tests. [Work supported by US Dept. of Agriculture.]

THURSDAY AFTERNOON, 19 MAY 2005

REGENCY F, 1:30 TO 4:45 P.M.

Session 4pNS

Noise, Architectural Acoustics and Engineering Acoustics: Characterization of Acoustical Materials

Brandon D. Tinianov, Chair

Quiet Solution Inc., 522 Almanor Ave., Sunnyvale, CA 94085

Invited Papers

1:30

4pNS1. On the measurement of the mechanical properties of acoustic porous materials. Nouredine Atalla (Dept. of Mech. Eng., Universit de Sherbrooke, Sherbrooke (QC), Canada J1K 2R1), Franck Sgard, and Luc Jaouen (Ecole Nationale des Travaux Publics de l'Etat, 69518 Vaulx-en-Velin Cedex, France)

This paper discusses the measurements and use of the mechanical properties of porous materials. It starts with a review of the classical methods currently used and concentrates in explaining the assumptions, difficulties and limitations of these methods. Next, it concentrates on two methods: a quasi-static approach accounting for the effect of shape factors on the measured properties and a hybrid inverse method similar to the vibrating beam method used for viscoelastic materials. Through full 3-D numerical simulations, both methods are simulated and their main features, assumptions, and limitations are investigated. In particular, the model uses full poro-elastic modeling and boundary element in order to investigate the effect of acoustic radiation damping and viscous damping. Using various types of materials, comparisons between these two methods and experimental data are also discussed.

1:50

4pNS2. Measurements of the acoustic properties of reticulated vitreous carbon. Ralph T. Muehleisen (Civil and Architectural Eng., Illinois Inst. of Technol., Chicago, IL 60616, muehleisen@iit.edu), C. Walter BeamerIV (Univ. of Colorado, Boulder, CO 80309), and Brandon D. Tinianov (Quiet Solution, Sunnyvale, CA 94085)

Reticulated vitreous carbon (RVC) is an open-cell foam structure composed of amorphous carbon that has a high porosity, strength and rigidity, low bulk thermal conductivity, a high melting point and chemical inertness. These properties make RVC an excellent candidate for use as a material for stacks and regenerators in thermoacoustic devices as well as an acoustic absorber in harsh environments. The four-microphone transfer matrix method was used to measure the characteristic impedance and wave number of 60 to 300 pore-per-inch RVC. From the measurements, a new empirical power law model and acoustic absorption charts were developed. The data are shown to be poorly predicted by the empirical models of Delany and Bazley and only fairly predicted by the Johnson-Allard microstructural model.

4pNS3. Prediction of the thermal conductivity of fiberglass insulation using propagation constant: A technique overview. Brandon Tinianov (Quiet Solution, Inc., 522 Almanor Ave., Sunnyvale, CA 94085), Masami Nakagawa, and David Muñoz (Colorado School of Mines, Golden, CO 80401)

This paper describes a novel technique for the prediction of the thermal performance of low density fiber glass insulation and other related fibrous insulation materials using a non-invasive acoustic apparatus. The project motivation is to create an enabling technology for in situ quality control testing of such fiber glass batts during production. Experimental results obtained in the laboratory show excellent correlation between the thermal conductivity and both the real and imaginary components of the propagation constant. Correlation of calculated propagation constant magnitude versus measured thermal conductivity gave an R^2 of 0.94 for the range of typically manufactured fiber glass batt materials. Given the promise of such highly correlated measurements, the acoustic technique could be used to continuously predict the thermal conductivity of the material during its production, replacing current off-line methods. The mechanisms for energy transfer through the materials are distinctly different for the acoustical propagation versus the heat transfer. The nature and behavior of the mechanisms will be reviewed and compared.

4pNS4. Estimation of the accuracy of measured acoustical parameters of porous materials with an acoustical method. Xavier Olny (ENTPE-DGCB URA, CNRS 1652, rue Maurice Audin 69518, Vaulx en Velin, France) and Raymond Panneton (Université de Sherbrooke, PQ, Canada J1K 2R1)

Acoustical parameters such as tortuosity, viscous and thermal characteristic lengths, or even static thermal permeability are often required to describe the dynamic behavior of porous materials. Using a middle frequency method, based on the use of a standing waves tube, the direct measurements of the resistivity and porosity, and the accurate measurement of the dynamic density and bulk modulus, these intrinsic parameters can be obtained from analytical inversion of Johnsons and Lafarges models. The interests of the method lie in the simplicity of the apparatus and on the clear separation of viscous and thermal dissipative contributions. This last point specially helps to understand the physics of the medium and check the validity of the assumptions made: the material should not have a strong elastic behavior, and it must fit the models used in the inversion process. In this presentation, the robustness of the method is discussed and estimations of the measurement uncertainties are given. The occurrence of systematic errors coming from the use of semi-phenomenological models for determining intrinsic parameters is tackled. The measurement procedure, and the experimental set-up are detailed, and results obtained on materials with very different properties are presented.

Contributed Papers

4pNS5. Model to simulate the acoustical absorption of laminated fiberglass insulation. Juan Arvelo, Jr. (Appl. Phys. Lab., Johns Hopkins Univ., 11100 Johns Hopkins Rd., Laurel, MD 20723-6099), Ilene Busch-Vishniac, James West, and Robert Ng (Johns Hopkins Univ., Baltimore, MD 21218-2681)

While fiberglass insulation is known to be a very good acoustical absorber in the audible frequency range, it is not often used for this purpose in rooms because of its appearance and its delicate structural rigidity. Additionally, in some situations, notably medical facilities, the material is inappropriate because of its potential to trap and retain bacteria and dirt. In this work, we report on a study of laminate-coated fiberglass. Although the addition of a laminate coating degrades the acoustical absorption, previous experimental work demonstrated that the composite structure retains a high absorption coefficient. A model of the wall with the installed fiberglass/laminate composite was developed and a wavenumber integration approach was applied to this model to assess the impact of the laminate on the performance of the system. The estimated reflection and transmission losses are computed as functions of auditory frequency and incidence angles with and without the laminate layer to compare the distributions of acoustical energy. This work is performed under a Hafstad fellowship sabbatical appointment.

4pNS6. Impact of acoustic leakage on the absorption of mono-layer and two-layers porous material. Franck Castel, Franck Sgard (Laboratoire des sciences de l'Habitat, DGCB URA CNRS 1652, Ecole Nationale des Travaux Publics de l'Etat, 69518 Vaulx-En-Velin Cedex, France), and Nouredine Atalla (Universite de Sherbrooke, Sherbrooke, QC, Canada J1K 2R1)

This paper discusses the effects of small lateral air gaps on the normal incidence absorption coefficient of mono layer and two layers porous materials. Such mounting conditions are responsible for changes in the absorption, leading to dramatic errors in the determination of the acoustics parameter with inverse characterization methods. As this type of mounting conditions is hard to control experimentally, a hybrid finite element-modal method is used to investigate the problem of the porous material inserted in a rectangular wave-guide. At the interface of the material and the wave-guide, coupling between the different domains is accounted for accurately using a modal decomposition. An automatic meshing approach is employed to speed up and guarantee convergence of the method. A large set of materials spanning a wide range of flow resistivities is used for the simulations. The results are presented under the form of charts which makes them an easy to use tool suitable for both inspection and design. Firstly, these charts allow one to identify materials whose normal inci-

dence absorption coefficient is sensitive to lateral air leaks. Secondly, these charts are a helpful tool for designing highly absorptive solutions based on the combination of porous materials and air gaps.

3:30

4pNS7. Sound absorbers with micro-perforated stretched foils and porous materials. Christian Nocke, Catja Hilge (Akustikbuero Oldenburg, Alte Raad 20a, D-26127 Oldenburg, Germany, info@akustikbuero.info), and Jean-Marc Scherrer (Barrisol Normalu SA, F-68680 Kembs, France)

At the spring meeting 2004 in New York first results for the sound absorption of micro-perforated stretched foils have been presented. The classical set-up of a micro-perforated sound absorber consists of a micro-perforated panel in front of an air cavity. The sound absorption coefficient of these set-ups can easily be calculated with a high accuracy according to the well-known approximation of D.-Y. Maa if all defining geometrical parameters (diameter of microperforation, distance between orifices, panel thickness and air cavity depth) are known. In this contribution measured sound absorption coefficients of other set-ups with micro-perforated foils as well as combinations with different porous materials will be presented. For these assemblies no closed calculation model exists so far. Comparisons between computer-based calculations of layered absorber set-ups show promising results for some assemblies. Finally different applications in various rooms will be presented. As expected the comparison between theoretical design and the final result show a very high degree of agreement.

3:45

4pNS8. Measured characteristic impedance and propagation constant of some common materials using the two cavity method. Richard Godfrey (Owens Corning, Sci. & Technol., 2790 Columbus Rd., Granville, OH 43023)

The acoustics of porous rigid frame materials is characterized by their characteristic impedance and propagation constant. These properties can be measured using the two cavity method. A sample is placed in an impedance tube with an air space behind the sample and ahead of the rigid termination. Measurements are made in accordance with ASTM E 1050. This procedure is repeated with the same sample, but a different air space. These data are analyzed as outlined by Seybert *et al.* [J. Acoust. Soc. Am. 86, 637 (1989)] to yield the characteristic impedance and propagation constant. This method was used to characterize fiber glass boards with a range of flow resistivities from 2600 to 56000 mks rays/m. Repeat measurements were made, and these measurements were compared with the Delaney and Bazley model, [Applied Acoustics (3), 1970]. Measurements were also made for cotton shoddy and cellulose, and power law models were developed. The method produced repeatable results comparable to the Delaney and Bazley model with little effort, and it is a useful method of obtaining property data for modeling tools such as Mechels Acoustics Program System (MAPS).

4:00

4pNS9. Estimation of the absorption coefficients of classroom surfaces by multiple regression. Katrina Scherebnyj and Murray Hodgson (UBC Acoust. & Noise Res. Group, SOEH, 3rd Fl., 2206 East Mall, Vancouver, BC, Canada V6T 1Z3)

The absorption coefficients of room surfaces are important parameters for predicting room sound fields. Of particular interest are classrooms, for which characteristics such as speech intelligibility are essential. Values

obtained in laboratory settings may differ significantly from those measured in real-world situations. This study determined surface absorption coefficients in university classrooms and compared them with values given by manufacturers and published literature. Various surface categories were established, and the total area of each was measured in 104 UBC classrooms, along with the early decay times (EDT) in octave bands from 125 to 8000 Hz when the rooms were unoccupied. Average octave-band absorption coefficients for these surfaces were determined from the measured EDTs, using diffuse-field theory. Multiple-regression analysis was then applied, and the data was fit to estimate the average absorption coefficients of each surface category in each octave band. Absorption coefficients were established for seven types of surfaces commonly found in university classrooms; in general they tended to be significantly lower than commonly accepted values.

4:15

4pNS10. Effects of damping materials on heavy-weight impact noise in reinforced concrete floor. Jin Yong Jeon, Young Jeong, and Seung Yup Yoo (School of Architectural Eng., Hanyang Univ., Seoul 133-791, Korea)

Damping materials for reducing heavy-weight floor impact noise in reinforced concrete structures were tested in apartment buildings. The effect of damping materials and an impact isolator were compared in an on-site experiment conducted in a high-rise apartment building. The results showed that the resonance frequency increased and vibration acceleration level decreased when the damping materials were used. Heavy-weight impact sound levels of the structure decreased substantially at 63 Hz, whereas the sound levels of the structure with the impact isolator increased. Subjective evaluation of the perceptual effects of the damping material and the impact isolator were also conducted in the laboratory.

4:30

4pNS11. In situ measurement of absorption of acoustic material with a parametric source in air. Volker Mellert and Roland Kruse (Oldenburg Univ., Inst. Phys., 26111 Oldenburg, Germany)

Measurement of properties of acoustic material in situ is quite desirable in numerous applications. But pulse-echo methods often lack from spurious reflections in a closed-space environment. It is therefore necessary to irradiate the sample material with a highly focussed acoustic beam of defined width. The principle of parametric transducers, which is well known in underwater acoustics, can be used in air as well, if the sound pressure of the ultrasound carrier is high enough. This was recently demonstrated by Hibrál and Zakharia (Proc. of the Joint Congr. of CFA/DAGA 04, Strasbourg 2004, p. 541). The new measurement set-up is in principle best suited for in-situ measurements due to the narrow beam-width. But ultrasound intensity, efficiency of generation of audio sound etc. give rise to several limitations. Applications of in-door measurements, in particular at grazing-angle, are presented.

Session 4pPAa**Physical Acoustics, Biomedical Ultrasound/Bioresponse to Vibration and ASA Committee on Standards: Cavitation and Other Mechanical Effects in Biomedical Ultrasound: A Special Session to Honor the Work of Wesley Nyborg II**

Lawrence A. Crum, Cochair

Applied Physics Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698

Junru Wu, Cochair

*Dept. of Physics, Univ. of Vermont, Burlington, VT 05405***Invited Papers****2:00****4pPAa1. How does ultrasound induce strain related responses in cells?** James Greenleaf and Mark Bolander (Mayo Clinic College of Medicine, 200 First St., Rochester MN 55905)

The fact that modulated ultrasound at a pulse intensity of 160 mW/cm^2 can induce fracture healing implies that cells that are sensitive to strain can feel and respond to this signal. A hypothesis that uses Nyborg inspired methods to investigate this effect will be presented along with data from a wide variety of simulations and experimental data. It appears that the response is mediated through the cytoskeleton and may be associated with interaction of the sound with focal adhesion complexes at one end of the skeletal fibers for high frequency responses and with the nuclear membrane at the other end of the fibers for low frequency responses. Although this hypothesis is nascent it is supported with experimental data from the literature.

2:20**4pPAa2. The influence of agent delivery mode on cardiomyocyte injury induced by myocardial contrast echocardiography in rats.** Douglas Miller, Chunyan Dou, and William Armstrong (Univ. of Michigan, Ann Arbor, MI 48109)

Bioeffects induced in rat hearts by the interaction of ultrasound with contrast agent gas bodies depends on user controlled parameters. This study examined the influence of contrast agent delivery mode on cardiomyocyte injury. Hairless rats were anesthetized and mounted vertically in a water bath. Evans blue dye was injected as stain for cardiomyocyte injury. The contrast agent (Definity, Amersham Health Inc., Princeton NJ) was diluted in saline and injected IV at 20 or 80 $\mu\text{L/kg}$ either as bolus or continuous infusion. Echocardiography was performed with a GE Vingmed System V in a short axis view with 1:4 or 1:16 ECG triggering at 1.5 MHz for 5 or 20 min. The peak rarefactional pressure amplitude (PRPA) was 2.0 MPa. Fluorescent stained cells were scored on frozen sections of heart samples obtained 24 hr after scanning. Five min infusion led to 3.4 times ($P < 0.01$) more cardiomyocyte injury than bolus injection for the same 20 $\mu\text{L/kg}$ dose. Contrast agent dose appeared to be a more important parameter than infusion rate and number of triggers. Contrast agent delivery mode, as well as dose and PRPA, has a significant influence on the bioeffects potential of myocardial contrast echocardiography. [Work supported by NIH grant EB00338.]

2:40**4pPAa3. Nonlinear model for detection of bubble activity from pulse-echo ultrasound.** Emad S. Ebbini (200 Union St. SE, Rm. 4-174, Minneapolis, MN 55455)

Cavitation and transient cavitation plays an important role in modern application of biomedical ultrasound, both in therapeutics and diagnostics. Special imaging and signal processing methods have been recently introduced for the detection and localization of cavitation and other nonlinear phenomena for improving image quality and/or characterization of therapeutic outcome of the ultrasound-based treatments. We have recently developed a model-based approach for separating the linear and nonlinear components of the pulse-echo data. In its simplest form, the approach separates the linear and quadratic signal components from the echo data. The latter is related to the sum and difference frequency interactions in the nonlinear medium. When used in conjunction with elongated coded transmit waveforms (e.g., chirp signals), the approach provides significant separation of bubble activity from tissue nonlinear components. Application to ultrasound contrast-agent imaging and detection of HIFU-induced lesions will be discussed.

3:00**4pPAa4. Effect of clinical diagnostic ultrasound pressure amplitude and pulse repetition frequency on echogenic liposome destruction.** Denise A. B. Smith, Tyrone M. Porter, Christy K. Holland (Dept. of Biomed. Eng., Univ. of Cincinnati, Medical Sci. Bldg, Rm. 6152, 231 Albert Sabin Way, Cincinnati, OH 45267-0586, smitdn@ email.uc.edu), Shao-Ling Huang, Robert C. MacDonald (Northwestern Univ., Evanston, IL 60208), and David D. McPherson (Northwestern Univ., Chicago, IL 60611)

Liposomes are submicron-sized phospholipid vehicles that contain both gas and fluid. Entrapped microbubbles within the liposome cause them to be echogenic. With antibody conjugation and thrombolytic drug incorporation, such as rt-PA, these liposomes can be used as novel targeted diagnostic ultrasound echo contrast agents to deliver a drug locally. In this study, the echogenicity of

unconjugated liposomes suspended in distilled water was assessed by measuring the decay rate of backscattered intensity using a Philips HDI5000 diagnostic ultrasound scanner with a 12-5 MHz linear array. Pulses of ultrasound with a 7-MHz center frequency and derated peak rarefaction pressure amplitudes of 0.16, 0.35, and 0.61 MPa were used to expose the liposomes. Pulse repetition frequencies in the range of 6.7 kHz to 13.3 kHz were employed at frame rates between 16 and 35 Hz, respectively. The average backscatter intensity was determined from a 20 mm² region-of-interest in each image frame. An exponential decay was fit to each data set. Ultrasound pulses with higher acoustic pressures and higher pulse repetition frequencies destroyed the liposomes at a faster rate.

3:20

4pPAa5. Cavitation detection during ultrasound-assisted thrombolysis in porcine blood clots. Saurabh Datta (Dept. of Biomed. Eng., Univ. of Cincinnati, Cincinnati, OH 45267), Louis E. McAdory (Univ. of Cincinnati), Jun Tan, and Christy K. Holland (Univ. of Cincinnati, Cincinnati, OH 45267-0586, Christy.Holland@uc.edu)

Substantial enhancement of recombinant tissue plasminogen activator (rt-PA) mediated thrombolysis can be achieved with exposure of a thrombus to pulsed ultrasound. However, the mechanism of this interaction has not yet been elucidated. In this work cavitation is investigated as a possible mechanism for enhancement in 120-kHz pulsed ultrasound-assisted thrombolysis. Porcine blood clots were immersed in plasma as control and exposed to rt-PA, ultrasound (0.35 MPa, 80% duty cycle, 1667 Hz pulse repetition frequency), or a combination of rt-PA and ultrasound. A confocally aligned active and passive cavitation detection system was employed to detect subharmonic emissions from stable cavitation and broadband superharmonic emissions from inertial cavitation. After exposure, clot mass loss was determined, and clots were subjected to immunohistochemical analysis of fibrin degradation. Spatial investigation of cavitation thresholds inside the clot, on the clot surface, and in the fluid surrounding the clot showed the threshold to be lowest on the clot surface. Stable cavitation was detected in clots exposed to ultrasound alone and a combination of rt-PA and ultrasound. Curiously, inertial cavitation was detected only in samples containing rt-PA. The presence of both stable and inertial cavitation correlated with increased clot mass loss and a distinct pattern of fibrin degradation on histologic evaluation.

3:40–4:00 Break

4:00

4pPAa6. Occlusive thrombosis in the rabbit auricular vein *in vivo* targeted by induction of intraluminal cavitation using HIFU and ultrasound contrast agent. Andrew A. Brayman, Juan Tu, Thomas Matula, Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698), Joo Ha Hwang, and Michael B. Kimmey (Univ. of Washington, Seattle, WA 98195)

Hypotheses tested: (1) inertial cavitation [IC] could be induced in the venous lumen *in vivo* by combined use of intravascular microbubble contrast agent and transcutaneous application of 1-MHz high intensity focused ultrasound [HIFU] of very low duty factor, and that IC activity could be detected and quantified *in vivo* as in earlier *in vitro* studies via passive cavitation detection methods; (2) robust IC activity would damage the venous endothelium in treated regions; (3) endothelial damage would be proportional to the IC dose developed in the region; (4) severe local endothelial damage alone may be sufficient to induce occlusive thrombosis, or may sensitize the region to low systemic doses of prothrombotic agents, and (5) biologically significant temperature rises and attendant thermal bioeffects in the vessel and perivascular tissues would not occur, even under the highest amplitude acoustic conditions applied. Each hypothesis was supported by the data. The principal result was that under treatment conditions involving very high peak negative acoustic pressures and contrast agent, treated areas thrombosed acutely but non-occlusively. When fibrinogen was administered locally after such treatment, occlusive thrombi formed acutely and only in the treated region, a response observed with none of the other treatments.

4:20

4pPAa7. Controlled nucleation of microcavitation with laser-illuminated nano-particles. Ronald A. Roy, Caleb H. Farny, Tianming Wu, Todd W. Murray (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215), and R. Glynn Holt (Boston Univ., Boston, MA 02215)

The safe utilization of controlled cavitation in HIFU therapy requires the presence of nucleation sites for bubble formation. Effective cavitation nuclei do not exist in most tissues; nucleation threshold pressures in excess of 4–5 MPa have been reported. We investigate the efficacy of transient vapor cavity generation from laser-illuminated gold nano-particles as a means for nucleating cavitation with high-intensity focused ultrasound. An acrylamide tissue phantom seeded with 82-nm diameter gold particle was exposed to 20 ns pulses from a 532 nm Nd:Yag laser. Laser firing was precisely synchronized with a pulsed 1.1 MHz HIFU pressure field. Acoustic emissions from inertial cavitation were detected by a 15 MHz focused transducer at a laser energy of 0.10 mJ/pulse and a HIFU peak-negative focal pressure as low as 0.92 MPa. In comparison, a peak-negative focal pressure of 4.50 MPa was required to nucleate detectable cavitation without laser illumination; nano-particles were present in both cases. Since the particles are durable, one can re-activate them as needed, essentially yielding cavitation nuclei on demand. [Work supported by the Dept. of the Army (award No. DAMD17-02-2-0014) and the Center for Subsurface Sensing and Imaging Systems (NSF ERC Award No. EEC-9986821).]

4:40

4pPAa8. Sonochemical synthesis and modification of protein microspheres. Kenneth S. Suslick, Farah J.-J. Toublan, and Elizabeth M. Dibbern (School of Chemical Sci., Univ. of Illinois at Urbana-Champaign, 600 S. Mathews Av., Urbana, IL 61801, ksuslick@uiuc.edu)

The need for organ-targeted delivery of drugs and imaging agents creates an interest in biocompatible, biodegradable vesicles. We make protein microspheres using high-intensity ultrasound; these microspheres have a protein shell and a hydrophobic interior making them ideal for delivering hydrophobic materials. We have previously shown that various proteins, e.g., bovine serum albumin (BSA), form a microsphere shell stabilized by inter-protein cross-linking of cysteine residues. In this study, we explore methods to modify the surface of these microspheres by (1) nanoparticle incorporation and (2) layer by layer electrostatic adhesion of macromolecules. We have explored applications of these microspheres for organ targeted imaging using both MRI and optical coherence tomography (OCT).

4:55

4pPAa9. Nonlinear dynamics of gas bubbles in soft tissue. Xinmai Yang and Charles Church (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677)

Understanding the behavior of cavitation bubbles driven by ultrasonic fields is an important problem in biomedical acoustics. The previous studies are largely limited by availability of experimental data on soft tissue. To approach this problem, we combine the Keller-Miksis equation for nonlinear bubble dynamics with the linear Voigt model for viscoelastic media. Using experimentally determined values of viscoelastic properties of soft tissue as a guide, the effects of elasticity on bubble oscillations are studied. The inertial cavitation thresholds are determined using a criterion of $R_{\max}/R_0=2$, and subharmonic emissions from an oscillating bubble are calculated. The results show that the presence of the elasticity increases the threshold pressure for inertial cavitation, and subharmonic signals only can be detected in a certain region of radii and driving pressures at a given frequency. These results should be useful in cavitation detection and bubble-enhanced imaging work. The model also could be used to determine values for the viscoelastic properties of soft tissue.

THURSDAY AFTERNOON, 19 MAY 2005

PLAZA A, 2:00 TO 4:15 P.M.

Session 4pPAb

Physical Acoustics: Computational Methods in Physical Acoustics

T. Douglas Mast, Chair

Univ. of Cincinnati, Biomedical Engineering, 231 Albert Sabin Way, Cincinnati, OH 45267-0586

Contributed Papers

2:00

4pPAb1. Acoustic Casimir pressure for arbitrary media. Raul Esquivel-Sirvent, Luis Reyes, and Jeffrey Barcenas (Instituto de Fisica, UNAM Apdo. Postal 20-364, Mexico DF 01000, Mexico, raul@fisica.unam.mx)

We derive a general expression for the acoustic Casimir pressure between two parallel slabs made of arbitrary materials and whose acoustic reflection coefficients are not equal. The formalism is based on the calculation of the local density of modes using a Green's function approach. The results for the Casimir acoustic pressure are generalized to a sphere/plate configuration using the proximity theorem.

2:15

4pPAb2. Use of complex source points to simplify numerical Gaussian beam synthesis. Stephen Forsythe (Naval Undersea Warfare Ctr., 641 Middle Rd., Portsmouth, RI 02871, stephen.forsythe@verizon.net)

It is often desirable to generate the acoustic field due to a so-called Gaussian beam. One way to do this is to use the free-space Greens function for the acoustic field and to sum small area sources over a circular plate with the appropriate shading for the desired Gaussian beam. For very high frequencies and narrow beams, the computation time to give an accurate sum can be large when calculating the sum for many points in the acoustic field. An alternate approach comes from the use of a single point source with complex coordinates $R=[Xr+iXi, Yr+iYi, Zr+iZi]$. When this complex source point is used in the free-space Greens function, the formal expressions for pressure and particle velocity can be used if careful attention is paid to the interpretation of the complex distance, r , that arises in the $\exp(ikr)/r$ term. The singularity is no longer a single point in the

case of a complex source, but a circular disk. The far field of a complex source point is a good approximation to a Gaussian beam. Several computational uses of the technique will be demonstrated. Extension to the shear wave Greens function will be explored.

2:30

4pPAb3. Causality and acoustic group velocity in dispersive media. Joel Mobley (U.S. Army Res. Lab., 2800 Powder Mill Rd., Adelphi, MD 20783, jmobley@arl.army.mil)

This talk examines two specific issues related to causality and the group velocity of acoustic pulses in dispersive media. First, the causal prediction of group velocity from attenuation with finite-bandwidth Kramers-Kronig ($K-K$) relations is discussed. Without extrapolating beyond the measurement bandwidth, the causal linkages of resonant-type data are established using expressions derived from the acoustic $K-K$ relations to predict group velocity and slope of attenuation (frequency derivative of attenuation). These predictions provide a stricter test of causal consistency than the determination of the phase velocity and attenuation coefficient due to their shape-invariance with respect to subtraction constants. Secondly, conditions under which the group velocity is the velocity of the peak in the envelope of a acoustic pulse are described. The spatial and temporal signatures of acoustic pulses with arbitrarily large (e.g., superluminal) and negative group velocities are also demonstrated. These signatures are shown to be consistent with both simple and relativistic causality.

4pPAb4. Ultrasonic backscatter in two-dimensional domains. Goutam Ghoshal and Joseph A Turner (Dept. of Eng. Mech., W317.4 Nebraska Hall, Univ. of Nebraska–Lincoln, Lincoln, NE 68588-0526)

The scattering of elastic waves in polycrystalline materials is relevant for ultrasonic materials characterization and nondestructive evaluation (NDE). Ultrasonic backscatter measurements are used widely to extract the microstructural parameters such as grain size and also to detect flaws in materials. Accurate interpretation of experimental data requires robust scattering models. Line transducers are often used for ultrasonic experiments such that an appropriate model for these two-dimensional problems is needed. Here, a theoretical expression for the temporal backscatter is derived for such domains under a single-scattering assumption. The result is given in terms of transducer and microstructural parameters. In addition, the problem is examined in terms of numerical simulations using Voronoi polycrystals that are discretized using finite elements in a plane-strain formulation. The material properties of the individual Voronoi cells are chosen according to appropriate material distributions. Such numerical models also allow scattering theories, including the one discussed here, to be examined for well-controlled microstructures. Example numerical results for materials with varying degrees of scattering that are of common interest are presented. The numerical results are compared with the theory developed with good agreement. These results are anticipated to impact ultrasonic NDE of polycrystalline media. [Work supported by US DOE.]

3:00

4pPAb5. Local dynamical effects in inhomogeneous elastic waveguides. Victor T. Grinchenko (Inst. of Hydromechanics of NAS of Ukraine, 8/4 Zhelyabov Str., 03680 Kiev, Ukraine)

The properties of normal waves in elastic waveguides are essentially different than properties of normal waves in acoustical waveguides. As the physical reason of the differences one can indicate the specific phenomena of reflection of elastic waves from a free surface. When an elastic wave reflects from the surface, the energy of longitudinal waves transforms to shear waves and vice versa. It is shown that the conversion of the wave motion types sets condition for generation of local disturbances in wave fields when normal waves are reflected from geometrical and physical inhomogeneities. The paper presents quantitative and qualitative analysis of the local effects for different cases. It was shown that such effects occur in a simple elastic waveguide. The specific features and physical meaning of the edge resonance effect in semi-infinite waveguides is presented. It was shown that for the waveguides with geometrical and physical inhomogeneous interesting phenomena of anomalous transparency of discontinuity can occur for any frequency bands. For other frequency bands the same discontinuity can be as an effective reflector of incident normal wave.

3:15

4pPAb6. Simplified series expansions for radiation from a baffled circular piston. T. Douglas Mast (Dept. of Biomed. Eng., Univ. of Cincinnati, Cincinnati, OH 45267-0586, doug.mast@uc.edu) and Feng Yu (Univ. of Cincinnati, Cincinnati, OH 45221-0070)

Computation of acoustic radiation from a baffled circular piston continues to be an active area of investigation, both as a canonical problem and because of numerous practical applications. For time-harmonic radiation, exact series expansions are an attractive approach because they do not require numerical integration or limiting approximations. Here, series expansions due to Hasegawa, Inoue, and Matsuzawa [J. Acoust. Soc. Am. **74**, 1044–1047 (1983); **75**, 1048–1051 (1984)] are shown to reduce to simpler expressions suitable for numerical computations. For the region $r \geq a$, where a is the piston radius and r is the distance from the piston center, an exact solution is given by a series of spherical Hankel functions and Legendre polynomials with explicit, closed-form, position-independent coefficients. For the paraxial region $w \leq a$, where w is the distance from the piston axis, a second exact series expansion is valid for all axial distances z and reduces to the known analytic solution for w

$= 0$. These two expansions allow the radiated field to be computed at any point, with rapid convergence except for points near the circle bounding the piston. Example numerical results illustrate application of this method to ultrasonic sources.

3:30

4pPAb7. Fast calculations of time-harmonic near field pressure distributions produced by triangular sources. Duo Chen and Robert McGough (Michigan State Univ., 2120 Eng. Bldg., East Lansing, MI 48824, mcgough@egr.msu.edu)

An analytical expression is derived for the near field pressure response to a sinusoidal excitation for a triangular piston. The analytical expression eliminates a numerical singularity from the impulse response and therefore achieves much more rapid convergence than the impulse response. This results in shorter computation times relative to the impulse response for a given peak error value. These fast-converging expressions are represented by three integrals, where each integral expression corresponds to an edge of the triangle. Pressure fields are evaluated within a grid that includes the face of the triangular source, and computation times are compared with the impulse response for given values of the peak error. For a specified peak error of 1%, the rapidly converging expressions are at least 50% faster than the impulse response for equilateral triangles with sides ranging from one-half wavelength to 8 wavelengths. For a specified peak error of 10%, the rapidly converging expressions are at least 120% faster than the impulse response when applied to the same group of equilateral triangles. Different reductions in computation time are achieved by other piston and grid geometries.

3:45

4pPAb8. Tangible interactive interface using acoustic time reversal process. Ros K. Ing, Nicolas Quieffin (Sensitive Object, Res. and Development Dept., 10 rue Vauquelin 75005 Paris, France), Stefan Catheline, and Mathias Fink (Univ. Paris VII, 75231 Paris cedex 05, France)

Time reversal in acoustic is a very efficient solution to focus sound back to its source in a wide range of material including reverberating media. It expresses the following physical properties: a wave still have the memory of its source location. The concept presented in this paper consists in detecting the acoustic waves in solid objects generated by a simple human touch. In a second step, the information related to the source location are extracted from a virtual time reversal experiment in the computer. Then, an action (turn on the light or a CD player) is associated to each location. Thus, the whole system transforms solid objects into interactive interfaces. Compared to the existing acoustic techniques, it presents the great advantage to be simple and easily applicable on inhomogeneous objects with any shape. The number of possible touch locations at the surface of objects is shown to be directly related to the mean wavelength of the detected acoustic wave.

4:00

4pPAb9. Telecommunication in a disordered environment with iterative time reversal. Gabriel Montaldo, Geoffroy Lerosey, Arnaud Derode, Arnaud Tourin (LOA, ESPCI, CNRS, 10 rue Vauquelin, 75005 Paris, France), Julien de Rosny, and Mathias Fink (CNRS, 75005 Paris, France)

Recent researches in the area of wave propagation in random media applied to telecommunications showed that, contrary to intuition, reverberation or scattering of waves in a disordered medium can actually help to increase the information transfer rate. The key element therein is the ability of a communication system to exploit independent channels of propagation. We present a method to transmit digital information through a highly scattering medium. It is based on iterations of a time-reversal process, and permits to focus short pulses, both spatially and temporally, from a base antenna to different users. This iterative technique is shown to be more efficient (lower inter-symbol interference and lower error rate)

than classical time-reversal communication, while being computationally light and stable. Experiments are presented: digital information is conveyed from 15 transmitters to 15 receivers by ultrasonic waves propagating through a highly scattering slab. From a theoretical point of view, the

iterative technique achieves the inverse filter of propagation in the subspace of non null singular values of the time reversal operator. We also investigate the influence of external additive noise, and show that the number of iterations can be optimized to give the lowest error rate.

THURSDAY AFTERNOON, 19 MAY 2005

REGENCY C, 1:00 TO 5:00 P.M.

Session 4pPP

Psychological and Physiological Acoustics: Localization, Binaural Hearing and Physiology (Poster Session)

Beverly A. Wright, Chair

Northwestern Univ., Dept. Communication Science and Disorders, Evanston, IL 60208

Contributed Papers

All posters will be on display from 1:00 p.m. to 5:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:00 p.m. to 3:00 p.m. and contributors of even-numbered papers will be at their posters from 3:00 p.m. to 5:00 p.m.

4pPP1. Auditory and visual performance in an open sound field. Kim F. Fluitt, Tomasz Letowski, and Timothy Mermagen (U.S. Army Res. Lab., HRED-VAP, Bldg 520, Aberdeen Proving Ground, MD 21005)

The ability to estimate a distance to a sound source in an open field is an important element of situational awareness on the battlefield, and is affected by many technical and environmental conditions. A limited body of knowledge regarding auditory perception of sources located over long distances makes it difficult to develop models predicting auditory behavior on the battlefield. Results of previous studies have shown that people greatly underestimate distances. However, it is not known if there is a connection between a visual estimation, verbally reporting an estimate, and auditory estimation. The purpose of the present study was to compare listeners' visual, verbal reporting, and auditory estimates to sound sources 25 to 200 m from the listening position. Data were collected for auditory as well as visual estimations of distance expressed verbally. Twenty-two subjects (men and women, ages 18–25) participated in the study. Six types of sounds were presented from five loud speakers in random order. Test results indicate that auditory and visual estimation errors increased with distance and that at some distances the visual estimation errors exceeded auditory estimations errors, but the differences were not statistically significant. Specific results will be presented.

4pPP2. Effects of different amounts of brief training and rest on the generalization of learning from interaural-level-difference to interaural-time-difference discrimination. Jeanette A. Ortiz and Beverly A. Wright (Dept. of Commun. Sci. and Disord. and Inst. for Neurosci., Northwestern Univ., 2240 Campus Dr., Evanston, IL 60208-3550)

Training-induced improvements on perceptual skills can be enhanced by increasing the amount of training and by resting between training and testing. However, how these two factors affect the generalization of learning from a briefly trained condition to an untrained one is unknown. Here, listeners were trained on an interaural-level-difference (ILD) discrimination condition (4-kHz tones), then were tested for generalization to an interaural-time-difference (ITD) discrimination condition (0.5-kHz tones). The amount of training and the time between training and testing differed across four groups. Listeners tested 10 hours after training had significantly lower ITD discrimination thresholds than naive listeners ($n=94$), regardless of whether training lasted for 20 min ($n=14$) or 2 h (n

$=11$). Thus, when there was a long time between training and testing, learning generalized from ILD to ITD discrimination regardless of the amount of training. In contrast, listeners tested immediately after training on ILD discrimination showed generalization to ITD discrimination with 20 min ($n=14$), but not 2 h ($n=14$), of training, suggesting that without rest between training and testing, longer training disrupted generalization. Hence, increased training interfered with generalization from ILD to ITD discrimination for short, but not long, delays between training and testing. [Work supported by NIH.]

4pPP3. The impact of helmet design on sound detection and localization. Angelique Scharine (U.S. Army Res. Lab.–HRED)

The shape and absorptiveness of a ballistic helmet changes the sound signal reaching the soldier's ear. Although the attenuation of current helmets is minimal, soldiers have reported removing their Personnel Armor System for Ground Troops (PASGT) helmet when it was necessary to determine sound source direction. The Advanced Combat Helmet (ACH) has reduced ear coverage and absorptive internal padding that decreases localization errors. In an effort to provide more integrated headgear, the U.S. Army is testing a new, Scorpion (R4) helmet that has rings that allow the insertion of earmuffs and a modified version without the rings (R4-R). Two studies were conducted to measure the effect of these four designs on sound detection and localization. None of the helmets differed significantly in the degree of attenuation nor did attenuation correlate with localization. However, localization performance varied greatly. The current ACH and the R4-R had the least impact on localization. The older, unpadded PASGT, caused high error rates despite having the lowest attenuation. Despite being otherwise similar to the R4-R, the R4 impaired localization just as much as the PASGT and its rings increased the attenuation by more than a decibel.

4pPP4. Principal components analysis interpolation of head related transfer functions using locally-chosen basis functions. Jacob W. Scarpaci and H. Steven Colburn (Hearing Res. Ctr. and Dept. of Biomed Eng., Boston Univ., 44 Cummington St., Boston, MA 02215)

Spatial interpolation between measured head related transfer functions (HRTFs) may be accomplished by using principal components analysis (Kistler and Wightman, 1992). The data set can be reduced by using only

the most important basis functions (BFs) when reconstructing the magnitude spectrum of the HRTFs. Relatively good performance can be achieved by using a small fraction of the BFs required for an error free representation. In this study a comparison is made between two methods for choosing which basis functions are included in the HRTF reconstruction for a given position. The usual approach is to choose the subset of BFs which have the largest impact (common variance) on the global set of measured HRTFs. This will be compared to a method in which the subset of BFs are chosen to have the largest impact on the measured HRTFs in the area local to the position being reconstructed. Advantages and disadvantages to the two methods are discussed. [Work supported by NIH DC00100.]

4pPP5. Hearing space: Identifying rooms by reflected sound. Lawrence D. Rosenblum and Ryan L. Robart (Univ. of California, Riverside, Riverside, CA 92521)

The acoustic engineering literature has shown that listeners are sensitive to the sound reflecting properties of spaces so as to judge the general dimensions and acoustic quality of the space [e.g., J. S. Bradley, R. D. Reich, and S. G. Norcross, *J. Acoust. Soc. Am.* **108**, 651–661 (2000)]. There is also speculation on the perceptual processes listeners use to apprehend properties of spaces [e.g., R. K. Clifton, R. L. Freyman, and J. Meo, *Percept. Psychophys.* **64**, 180–188 (2003)]. Still, there has been little research testing whether listeners can simply recognize different types of rooms based on reflected structure. To examine this issue, five different sound sources were binaurally recorded in four different acoustic spaces. The sources ranged from human speech to white noise bursts. The acoustic spaces included a large indoor gymnasium, a moderately-sized classroom, a public restroom, and a small laboratory room. Untrained listeners heard each sound over headphones as they looked at photographs of the four rooms. They were asked to choose the room in which each sound was recorded. Overall, listeners were quite good at the task, with accuracy dependent on sound source. The results indicate a sensitivity to complex ambient structure not often addressed in the psychoacoustics literature.

4pPP6. Effects of various reverberant conditions on speech intelligibility. Suzanne P. Carr and H. Steven Colburn (Hearing Res. Ctr. and Dept. of Biomed. Eng., Boston Univ., 44 Cummington St., Boston, MA 02215)

Understanding speech in complex environments is a problem for many hearing impaired listeners, yet the specific characteristics of the listening environment which contribute most to listening difficulties are not well understood. Using virtual stimuli, a large range of conditions are tested on a small group of listeners, both normal and hearing impaired, and speech intelligibility thresholds are obtained. Systematic comparisons of the different conditions are made. The conditions that are varied are the amount of reverberation in the virtual room (ranging from pseudo-anechoic to moderately reverberant), the presence or absence of strong, early reflections such as from a wall and/or tabletop, the direction of the target speech relative to the listener, the distance of the target speech from the listener, and the number and location of competing speech maskers. Care is taken to ensure that the conditions are within the bounds of realistic listening. The validity of the use of virtual stimuli is also explored. [Work supported by NIH DC00100.]

4pPP7. The effect of spatial configuration in a divided attention task. Virginia Best, Antje Ihlefeld, and Barbara Shinn-Cunningham (Hearing Res. Ctr., Boston Univ., 677 Beacon St., Boston, MA 02215, ginbest@cns.bu.edu)

The effect of spatial separation on the ability of listeners to report keywords from two simultaneous talkers was examined. The talkers were presented with equal intensity at a clearly audible level, but were processed to minimize their spectral overlap and reduce energetic interfer-

ence. The two talkers were presented with various angular separations around references of -45° , 0° , or 45° azimuth. Overall, performance did not vary dramatically with spatial configuration, but depended on spatial separation and reference direction. With the talkers in front or to the left, performance tended to first increase and then decrease with increasing separation. With both talkers to the right, performance tended to improve monotonically with increasing separation. The relative levels of the two talkers at each ear in each configuration partially accounted for results. For each talker a different ear contained a signal-to-noise ratio advantage, and performance was positively correlated with the mean signal-to-noise ratio across the two “better ears.” Thus, when tracking two sources simultaneously, listeners may make use of the information at the two ears independently. Furthermore, the drop in performance for some large talker separations may reflect increased difficulty in following sources that do not fall within a single “spotlight” of spatial attention.

4pPP8. Localization of amplitude and frequency modulated sounds. Mark Ericson (Air Force Res. Lab., AFRL/HECB Bldg. 441, 2610 Seventh St., Wright-Patterson AFB, OH 45433, mark.ericson@wpafb.af.mil)

Everyday sounds, such as speech, music, and environmental noises, vary in level and frequency. The dynamic cues of natural sounds help to identify and segregate multiple sources. Amplitude and frequency modulation of simple pure tone complexes were manipulated in several experiments to measure their effects on auditory localization. The tonal complexes had 500 Hz fundamental frequencies and 26 components, providing 13 kHz in bandwidth. The complexes were either amplitude modulated, frequency modulated, co-modulated in frequency and amplitude, or not modulated. In addition, phase of the carrier and modulation frequencies were randomized to aid in sound source segregation. Three normal hearing subjects participated in the task of localizing and identifying up to four simultaneous sounds. Amplitude modulated sounds provided the best localization acuity and identification data. Comodulated, frequency modulated and no modulation were found to provide lower amounts of acuity performance in decreasing order.

4pPP9. Different interaural level difference processing with complex sounds and pure tones. Yuxuan Zhang and Beverly A. Wright (Dept. of Commun. Sci. and Disord. and Inst. for Neurosci., 2240 Campus Dr., Northwestern Univ., Evanston, IL 60208-3550)

Interaural level differences (ILDs) are one of the primary cues to sound source position on the horizontal plane. Most studies on human performance with ILDs use simple sounds such as pure tones. However, naturally occurring acoustic stimuli usually have complex waveforms. Here human learning of ILD discrimination was examined with complex waveforms. Sixteen listeners were trained 1 hr/day for 9 days on ILD discrimination with a 4-kHz tone sinusoidally amplitude modulated at 0.3 kHz conveyed through headphones. Before and after training, they were tested, together with sixteen untrained controls, on the trained condition and five related untrained conditions (three with amplitude modulated tones, two with pure tones). The trained listeners improved significantly more than controls on all conditions with amplitude-modulated tones, but not on those with pure tones. The lack of generalization of ILD learning from modulated to pure tones suggests that practice modified ILD processing in a region that encodes complex sounds with a variety of carrier frequencies and modulation rates, but not pure tones. Thus, these findings suggest that waveform complexity is an important factor in ILD processing, and must be considered when evaluating human ILD performance. [Work supported by NIH.]

4pPP10. Sources of variation in masking by competing speech.

Richard L. Freyman, Karen S. Helfer, and Uma Balakrishnan (Dept. of Commun. Disord., Univ. of Massachusetts, Rm. 3 Arnold House, Amherst, MA 01003, rlf@comdis.umass.edu)

The literature on speech recognition within a competing speech environment shows great variation with respect to the amount of masking produced, as well as variation in the improvement realized from spatial separation of the target and masking speech. Presumably, the sources of variation include differences in methodology and stimuli used by different researchers. In the current study the target speech stimuli were nonsense sentences spoken by a female talker. The competing speech was created from similar nonsense sentences recorded by 10 other female speakers. Five different maskers were created, each a combination of two of the 10 talkers who had similar fundamental frequencies. The results showed that the variation in masking produced by the different two-talker combinations was substantial only when target and masker were presented from the same loudspeaker. As a consequence, the observed benefit from spatial separation varied widely among the different maskers, and was ordered according to fundamental frequency in an unexpected way. The results are interpreted as showing greater variation in informational masking than in energetic masking. Also tested was the influence of different ways of blocking the trials, which effectively manipulated the uncertainty of the masking stimulus from trial to trial. [Work supported by NIH DC01625.]

4pPP11. Binaural informational masking release in children and adults.

Joseph W. HallIII, Emily Buss, and John H. Grose (Univ. of North Carolina at Chapel Hill, Chapel Hill, NC 27599, jwh@med.unc.edu)

Informational masking associated with masker spectral uncertainty can sometimes be reduced by providing cues that promote perceptual segregation of the signal from the masker. Our previous research using the multi-burst-same (MBS) paradigm of Kidd *et al.* [J. Acoust. Soc. Am. **95**, 3475–3480 (1994)] indicated that all adults and most children (5–9 YO) showed reduced informational masking when monaural segregation cues related to gating asynchrony or spectro-temporal coherence were available. The present study examined developmental effects in children aged 6–10 years for release from informational masking in a condition where perceptual segregation was manipulated using a spatial hearing cue. In the informational masking condition, the pure-tone signal and the MBS masker components were presented to the left ear. In the masking release condition, the masker components were also presented to the right ear synchronously (at a level 10-dB higher than in the left ear) so that the signal would be lateralized to the left ear and the masker would be lateralized to the right ear. Results indicated that most adults achieved masking release in this condition, but that none of the children did. These results are consistent with previous findings indicating protracted development for informational masking release related to spatial hearing cues.

4pPP12. Measuring the binaural temporal window.

Andrew J. Kolarik and John F. Culling (School of Psych., Cardiff Univ., Tower Bldg., Park Pl. Cardiff, CF10 3AT, UK)

Sensitivity to small ITDs in detection and discrimination tasks was measured. In experiment 1, four listeners performed 3 tasks. The first 2 were 4I- and 2I-2AFC adaptive detection tasks. The third was a 2I-2AFC adaptive discrimination task. All stimuli were 100 ms noise bursts. Reference stimuli were diotic, and target stimuli contained a probe duration of 64, 32, 16, 8, 4, or 2 ms, which carried an interaural time delay. These probes were temporally fringed with diotic noise. Thresholds for the discrimination task were significantly higher than for the detection tasks. In experiment 2, psychometric functions were obtained from 4 participants for the six probe durations using the same 2I-2AFC detection and discrimination tasks. Binaural temporal windows were fitted to the data using a variety of fitting functions. Fits to the detection task data demonstrated narrow tips but unmeasurably long skirts. In the discrimination task neither parameter could be accurately measured, suggesting that the overall stimulus duration was too short to encompass the window. A stimulus

length of 500 ms and probe durations of 256, 128, 64, 32 and 16 ms allowed both parameters to be measured. The resulting equivalent rectangular duration was approximately 50 ms.

4pPP13. Individual differences in the masking level difference (MLD) with a narrowband masker at 500 or 2000 Hz.

Emily Buss, Joseph W. HallIII, and John H. Grose (Univ. of North Carolina at Chapel Hill, 130 Mason Farm Rd., 1115 Binformatics Bldg., CB7070, Chapel Hill, NC 27599, ebuss@med.unc.edu)

The MLD for a narrowband masker is associated with marked individual differences. This study examines factors that might account for these individual differences, including binaural temporal resolution and sensitivity to interaural cues based on time or level. MLD data were collected for 50-Hz wide maskers at 500 and 2000 Hz, gated on for 400-ms. The signal was a pure tone at the center frequency of the masker, of either brief (15-ms) or long (200-ms) duration. Brief signals were coincident with either a local minimum or maximum in the pattern of inherent amplitude modulation. Sensitivity to interaural time and level cues was assessed using similar stimuli, with the exception that the long-duration signal was a 50-Hz band of noise, generated to provide just interaural time cues, just interaural level cues, or both types of cues. Binaural temporal windows were estimated based on pure tone detection thresholds for a brief 500-Hz tone presented at different points in time relative to an abrupt interaural phase transition in a masking noise, bandpass filtered between 100 and 2000 Hz. Relationships between MLD results and those of binaural temporal windows and interaural cues of time and level will be discussed.

4pPP14. The discrimination of interaurally correlated noise bands.

Barrie Edmonds and John Culling (School of Psych., Tower Bldg., Cardiff Univ., Cardiff, CF10 3AT, UK)

Models of binaural unmasking suggest that the auditory system detects deviations from unity in the interaural coherence of the waveforms at the two ears. Since anticorrelated noise has high coherence within a frequency channel, such a coherence-detection mechanism should be unable to distinguish a correlation of -1.0 from a high positive correlation. A 3I-2AFC paradigm was used to test whether such a limitation can be observed behaviorally. Stimuli consisted of a 1-ERB-wide sub-band of interaurally correlated noise centered at 500 Hz (i.e. 462–538 Hz) flanked by two spectrally remote bands of noise (0–329 Hz and 717–3000 Hz) with an interaural correlation of 1.0. The flanking bands were intended to reduce cues from the perceived laterality of the stimuli. Participants were presented with three stimuli (reference, target, and distracter) and asked to choose the odd-one-out: the reference and distracter intervals contained sub-bands of anticorrelated noise while the target interval contained a sub-band of noise with interaural correlation between 0.7 and 1.0. Preliminary results indicate that different participants find different values of interaural correlation to be more distinguishable than others.

4pPP15. A biologically inspired binaural approach to monaural modeling.

Daniel E. Shub and H. Steven Colburn (Boston Univ. Hearing Res. Ctr., 44 Cummington St., Boston, MA 02215)

The auditory system is often discussed as having monaural and binaural neurological pathways; similarly models are classified as either monaural or binaural. Psychophysical evidence of contra-aural interference (when performance with one ear is better than performance with two ears) suggests that the information used on monaural tasks (e.g., N_0S_0 and N_mS_m detection) may be carried by a binaural pathway. Binaural models often require monaural channels to predict the results of monaural tasks, but these monaural channels prevent the models from predicting contra-aural interference. This modeling work investigates the monaural information carried by a processor which is inherently binaural. The processor design makes the inclusion of monaural channels unnecessary and contra-

aural interference is predicted under certain conditions. The performance of the model matches results from a variety of traditional psychophysical tasks (including discrimination of differences in overall intensity; discrimination of differences in interaural level, time and coherence; as well as detection under monaural and binaural masking conditions). Results suggest that binaural neurons contain sufficient information to explain performance on both binaural and monaural tasks. [Work supported by NIH grants R01 DC 00100 and 1 F31 DC006769-01.]

4pPP16. Auditory nerve response to broadband noise with high-frequency spectral notches. Ana Alves-Pinto, Enrique A. Lopez-Poveda (Instituto de Neurociencias de Castilla y Leon, Univ. of Salamanca, Avda. Alfonso X El Sabio s/n, 37007 Salamanca, Spain), and Alan R. Palmer (MRC Inst. of Hearing Res., Nottingham NG7 2RD, UK)

The threshold notch depth for discriminating between a flat-spectrum broadband noise and a similar noise with a rectangular spectral notch centered at 8 kHz varies nonmonotonically with stimulus level (Alves-Pinto and Lopez-Poveda, submitted to J. Acoust. Soc. Am.). A possible explanation for this result is that the notch may be encoded in the rate profile of auditory nerve (AN) fibers with high spontaneous rate (HSR) at low levels and in that of low spontaneous rate (LSR) fibers at high levels. To test this hypothesis, the rate profile of guinea pig AN fibers was measured in response to broadband notched noise for different notch depths and widths and for overall levels ranging from 40 to 100 dB SPL. Preliminary results support the hypothesis that HSR fibers can encode the spectral notch only for levels up to around 70 dB SPL. However, they also suggest that, as in the psychophysics, the negative effect of level is less pronounced the wider the notch. Additional data are required to confirm the role of LSR fibers to encode for the notch at higher levels. [Work supported by Spanish FIS PI020343 and G03/203.]

4pPP17. The spatial distribution of sound pressure within the human ear canal. Michael R. Stinson and Gilles A. Daigle (Inst. for Microstructural Sci., Natl. Res. Council, Ottawa, ON, Canada K1A 0R6, mike.stinson@nrc-cnrc.gc.ca)

The sound field inside a human ear canal has been computed using 2 approaches. A simple model, a modified Webster horn equation approach, can accommodate the curvature and varying cross section of the ear canal. Calculations using the horn equation demonstrate the formation of standing waves within the canal. The pressure may be interpreted as either the pressure along the center axis of the canal or the average pressure within a cross-sectional slab. To investigate possible spatial variation through a cross section, the sound field has also been computed using the boundary element method (BEM). Over 2000 triangular mesh elements, 1 mm or less in size, were used to represent the canal geometry. For a plane piston source at the canal entrance and both a rigid and a resistive impedance condition at the eardrum position, the computed sound pressures along the center axis of the ear canal are in good agreement with the horn equation calculations, up to 15 kHz. The BEM approach, though, reveals spatial variations of sound pressure through each canal cross section, increasingly significant as frequency increases. Further, for source configurations that are more realistic than a simple piston, large transverse variations in sound pressure are anticipated.

4pPP18. Distortion product otoacoustic emissions generated by mistuned harmonic stimuli. Glenis R. Long and Jungmee Lee (Speech and Hearing Program, Grad. Ctr. CUNY, 365 Fifth Ave, New York, NY 10016)

Psychoacoustic research in humans [Lee and Green, J. Acoust. Soc. Am. **95**, 718–725 (1994)] suggests that the detection of the mistuning of an harmonic when the harmonic complex and a mistuned component were presented simultaneously to the same ear stems, at least in part, from the resulting envelope interactions on the basilar membrane. Neurophysiologi-

cal research in chinchillas [Sinex *et al.*, Hear. Res. **168**, 150–162 (2002)], provides evidence that such envelope interactions can be detected at the level of the inferior colliculus. The potential role of cochlear nonlinearities in determining the nature of the signal on the basilar membrane is explored by evaluating the DPOAE generated by multicomponent harmonic and inharmonic complexes. When harmonic complexes are used, the DPOAE all fall at harmonic frequencies. When inharmonic complexes are used, many nonharmonic DPOAE are detected.

4pPP19. The reflectivity function of the cochlear transmission line. Renata Sisto (ISPESL–DIL, Via di Fontana Candida, 1, 00040 Monteporzio Catone (RM), Italy) and Arturo Moleti (Univ. of Roma Tor Vergata, 1, 00133 Roma, Italy)

The generation of evoked otoacoustic emissions is related, for each frequency, to the reflectivity function of the cochlear membrane near the correspondent resonant place. A formal analogy exists between the transmission line equation that is often used to model the propagation of sound along the cochlea and the Schrodinger equation describing the motion of an elementary particle in a one-dimensional potential. In this analogy, the cochlear region near the tonotopic place is equivalent to a negative potential well. Analytical solutions are available in quantum mechanics textbooks, which predict partial reflection of the incoming wave from such a potential well. The reflectivity is expected to be a slow quasi-periodic function of the well width and depth, e.g., of the quality factor of the cochlear resonance. This observation could be useful to predict the large-scale spectral structure of evoked otoacoustic emissions as a function of the cochlear tuning curve.

4pPP20. Multiple auditory steady-state response thresholds to bone-conduction stimuli using three oscillator placements in premature infants. Jennifer Hatton, Susan A. Small, and David R. Stapells (School of Audiol. & Speech Sci., Univ. of British Columbia, 5804 Fairview Ave, Vancouver, BC, Canada V6T 1Z3, stapells@audiospeech.ubc.ca)

Auditory Steady State Responses (ASSRs) are a promising technique for estimating the behavioral audiogram in infants. The current study investigates the effects of bone-oscillator placement on ASSR thresholds in premature infants ($N=15$, mean age: 35 weeks PCA). Using the MASTER research system, ASSR thresholds to multiple bone-conduction tones (0–50 dBHL re:mastoid; 77–101 Hz) presented to the (i) superior-posterior temporal bone, (ii) mastoid, and (iii) forehead were obtained. Results with no response at 50 dBHL were arbitrarily assigned a threshold of 60 dBHL. No differences in threshold were found between the temporal and mastoid placements. Thresholds averaged across these placements were 17, 15, 34, and 30 dBHL for 500, 1000, 2000, and 4000 Hz, respectively. However, on average, thresholds for the forehead placement were significantly higher by at least 14, 11, 18, and 14 dB at 500, 1000, 2000, and 4000 Hz, respectively, compared to the temporal and mastoid placements. These differences may be even greater because absent responses at the maximum intensity were seen more often with forehead placement compared to other placements. In conclusion, one may use either the temporal bone or mastoid for bone-oscillator placement, but the forehead should be avoided. [Work supported by CIHR and NSERC-Canada.]

4pPP21. Multiple auditory steady-state response (ASSR) thresholds to bone-conduction stimuli in young infants with normal hearing. Susan A. Small, Jennifer Hatton, and David R. Stapells (School of Audiol. & Speech Sci., Univ. of BC, 5804 Fairview Ave., Vancouver, BC, Canada V6T 1Z3)

Bone-conduction ASSR thresholds (carrier frequencies: 500–4000 Hz; 77–101-Hz modulation rates; amplitude/frequency modulated; single-polarity stimulus) were obtained in two infant groups ($N=29$ pre-term, tested in NICU; $N=14$ 0-to-8 months of age, tested in soundbooth). Mean (1SD) ASSR thresholds were 16 (11), 16 (10), 37 (10), and 33 (13) dBHL

in pre-term infants and 14 (13), 2 (7), 26 (6), and 22 (8) dBHL in older infants at 500, 1000, 2000, and 4000 Hz, respectively. Both infant groups had significantly better thresholds for 500 and 1000 Hz compared to 2000 and 4000 Hz, in contrast to adults who have similar thresholds across frequency (22, 26, 18, and 18 dBHL) [Small and Stapells, *J. Am. Acad. Audiol.* (in press)]. When 500- and 1000-Hz thresholds were pooled, NICU infants and 0–8-month-old infants tended to have better low-frequency thresholds than adults. When 2000- and 4000-Hz thresholds were pooled, NICU and 0–8 month-old infants tended to have poorer thresholds than adults. Overall, these results suggest that low-frequency bone-conduction thresholds worsen and high-frequency bone-conduction thresholds improve with maturation. [Work supported by CIHR and NSERC Canada.]

4pPP22. Human electrophysiological examination of the buildup of the precedence effect (Clifton effect). Andrew Dimitrijevic and David R. Stapells (School of Audiol. & Speech Sci., Univ. of British Columbia, 5804 Fairview Ave., Vancouver, BC, Canada V6T 1Z3, andrew@audiospeech.ubc.ca)

The relationship between behavioral measures of buildup of precedence effect [or Clifton effect (CE)] and electrophysiological responses using an event-related potential (ERP) paradigm was examined in 14 young adults with normal hearing. The CE was elicited using a binaural paired-click (left and right speaker delay) in sound field. This study aimed to determine whether ERP measures were related to the perception of buildup. Subjects participated in two sessions: (i) Session #1 psychoacoustic measures of the degree of buildup as function of click delay (left and right side). (ii) Session #2 recorded cortical ERPs using click delays (varying from 6 to 8 ms) that would maximize the degree of buildup in each subject. Stimuli consisted of repeated trains containing 5 paired-clicks (left speaker leading) with a noise burst separating consecutive trains. Results: Significant *N1* amplitude hemispheric asymmetries were observed related to click position. *N1* responses decreased (43%) as a function click position in the train at electrode sites contralateral to the leading stimulus and, conversely, increased (30%) at electrode sites ipsilateral to the lead. These results suggest that differential refractoriness of the *N1* response may lead to the lateralized perception of buildup to the precedence effect. [Work supported by NSERC-Canada.]

4pPP23. The event-related brain potentials to uncorrelated fragment of noise. Juan Huang, Xihong Wu, Liang Li (Dept. of Psych., Peking Univ., Beijing 100871, China, jacee@pku.edu.cn), James Qi, Bruce Scheneider (Univ. of Toronto at Mississauga, Mississauga, ON, Canada), Yu He, and Claude Alain (Baycrest Ctr. for Geriatric Care, Toronto, ON, Canada)

Two correlated waveforms of broadband noise presented to the left and right ear simultaneously or with a short delay are often perceptually fused into one sound image. However, when an uncorrelated noise fragment (UCNF) is inserted in a long duration sound, listeners report hearing a transient burst of noise. The detection of the UCNF is dependent on the duration of the UCNF and the binaural delays. Here, we recorded event-related brain potentials (ERPs) to 100-ms UCNF and varied the binaural delay (0, 3, 10, and 20 ms) from trial to trial. The likelihood of detecting the UCNF decreased with increasing binaural delay. At 0 delay, the UCNF elicited negative and positive waves peaking at about 100 and 200 ms after UCNF onset (N1-P2 complex), which was present even when the stimuli were ignored. The conscious detection of the UCNF elicited an additional positive wave between 250 and 450 ms at parietal and occipital sites (P3b). The P3b latency was longer and its amplitude larger for binaural delay of 3 than 0 ms. Our results show that the detection of UCF involved both automatic and attention-dependent processes especially when a binaural delay is introduced between two source of correlated noises.

4pPP24. Tinnitus neural map; A positron emission tomography study. Martin L. Lenhardt (Dept. of Biomed. Eng., Otolaryngol., Emergency Medicine, Virginia Commonwealth Univ., Richmond, VA 23298), Abraham Shulman, and Barbara Goldstein (SUNY, Brooklyn, NY)

The neural map of tinnitus involves more than the classical auditory pathways and the limbic system. PET findings in six patients, with severe problem tinnitus, revealed the involvement of the cerebellum, insula and frontal cortices when these patients were imaged before and after high frequency bone conduction therapy. The frontal cortex and cerebellum demonstrated the highest ratios of metabolic change but changes were also noted in the thalamus and the medial temporal lobe system. The PET data supports the view that the frequency specific map of auditory cortex is dynamic and can change with high frequency therapy, presumably due to neural reorganization. The PET data further reflect alterations in multiple areas of brain in all patients to tinnitus and/or high frequency therapy. Encouraging was the observation that patients with the most dramatic change in their global pre/post PETs were also the ones who exhibited the greatest behavioral improvement in regards to tinnitus relief measured by informal report and standard questionnaires. In those patients, post therapy minimal masking levels dropped indicating less acoustic energy was need to mask their problem tinnitus, interpreted as reflecting central changes. Taken as a whole, these data support the value of frequency specific tinnitus therapy.

4pPP25. Real-time implementation of a model auditory-nerve fiber. Satish G. Iyengar (Inst. for Sensory Res. and Dept. of Elec. Eng. and Comput. Sci., Syracuse Univ., Syracuse, NY 13244), Jayant Datta (Syracuse Univ. and Discrete Labs., Inc., Syracuse, NY 13220), and Laurel H. Carney (Syracuse Univ., Syracuse, NY 13244)

A time-varying model auditory-nerve (AN) fiber based on Tan and Carney (2003) was developed on a digital signal processor (DSP) platform to simulate AN responses in real-time, using the Motorola DSP56367, a 150 MIPS audio processor. To understand the schemes employed by the brain for decoding different sound stimulus parameters, it is first essential to get an accurate representation of peripheral (external, middle, and inner ear) responses that in turn excite the higher levels along the auditory pathway. Though present computational models predict essential AN response properties, their use in testing mechanisms hypothesized to occur at higher levels of the auditory pathway is limited due to the large processing times that they require. This is because such nonlinear models require computational implementation in the time domain, with time-step-by-time-step changes in model parameters as the bandwidth and gain of the inner ear change with stimulus fluctuations. The DSP-based model presented here aims to speed up inquiries related to auditory signal processing strategies at higher levels by providing a real-time solution to AN fiber simulations. Details of the implementation and some prospective auditory-research utilities will be discussed. [Work supported by NIDCD R43-DC006591 (JD, SGI) and R01-DC001641 (LHC, SGI).]

4pPP26. Modeling inner hair cell compression. Almudena Eustaquio-Martin and Enrique A. Lopez-Poveda (Instituto de Neurociencias de Castilla y Leon, Universidad de Salamanca, Avda. Alfonso X El Sabio s/n, 37007 Salamanca, Spain)

The voltage-dependent activation of inner hair cell (IHC) basolateral potassium (K^+) currents results in a compression of voltage responses to injections of large depolarizing currents with values well within the range of IHC transducer currents (Kros and Crawford, *J. Physiol.* **421**, 263–291 (1990)). *In vivo*, this compression must add to the compressive nonlinearity produced by the gating of transducer channels. The purpose of this work is to present a biophysical model of the IHC that simulates the nonlinear compressive effects of IHC basolateral potassium currents both *in vivo* and *in vitro*. The model incorporates two components for the po-

tassium currents, one fast and one slow, as described by Kros and Crawford. Membrane parameters are provided for the model to reproduce the time course of the receptor potential in response to constant current pulses across the membrane. The model is then used to investigate the degree of

compression that the K⁺ currents contribute to the receptor potential *in vivo*. Also investigated are the effects on IHC compression of a partial blocking of the K⁺ channels, as occurs in some common forms genetic hearing loss. [Work supported by Spanish FIS PI020343 and G03/203.]

THURSDAY AFTERNOON, 19 MAY 2005

PLAZA B, 2:00 TO 4:45 P.M.

Session 4pSA

Structural Acoustics and Vibration and Musical Acoustics: Vibration of and Acoustic Radiation from Musical Instruments II

Courtney B. Burroughs, Cochair

Applied Research Lab., Pennsylvania State Univ., State College, PA 16804-0030

Thomas D. Rossing, Cochair

Physics Dept., Northern Illinois Univ., De Kalb, IL 60115

Invited Papers

2:00

4pSA1. Piano soundboards: Experimental and modeling studies. Nicholas Giordano (Dept. of Phys., Purdue Univ., West Lafayette, IN 47907, giordano@purdue.edu)

The piano has attracted the attention of many researchers, and most of that work has focused on piano hammers and string vibrations. The vibrational and radiative properties of the soundboard have drawn much less attention, but are essential for an understanding of the sound produced by the instrument. We review experimental studies of the soundboard's vibrational modes and their radiative efficiency. We also describe modeling studies of soundboard vibrations and radiation, which aid in understanding the measurements. Prospects for future work and for applications to other instruments are suggested. [Work supported by NSF grant PHY-9988562.]

2:30

4pSA2. Phase mapping and sound radiation from Caribbean steelpan. Andrew Morrison (Dept. of Phys., Northern Illinois Univ., DeKalb, IL 60115)

The Caribbean steelpan is one of the most interesting acoustic musical instruments invented in the last century. Although simple in design, the acoustic properties of the steelpan are surprisingly complicated. Holographic interferometry was used to determine the resonances of a low tenor steelpan. Placement of a vibrating mirror in the optical path of the reference beam expands the capabilities of the holography system to include phase measurements. Phase maps of several low resonances of notes on a Caribbean steelpan will be shown. Sound intensity measurements were taken to explore the relationship between the resonances and the radiated sound field. The pan was placed in an anechoic chamber, and selected notes were excited electromagnetically with a swept sinusoid signal. A two-microphone probe was used to gather sound intensity measurements. Sound intensity maps of the first three harmonics will be shown for several notes.

3:00

4pSA3. Acoustics of Baltic psaltery; a new outstanding Latvian Kokle. Andres Peekna (Innovative Mech., Inc., 5908 N River Bay Rd., Waterford, WI 53185) and Thomas D. Rossing (Northern Illinois Univ., DeKalb, IL 60115)

We have studied the acoustics of a Latvian Baltic Psaltery (kokle) which was judged by performing and recording musicians to be outstanding. Previous work by the authors pointed out the importance of a high population of body-resonances within the tuning range of the instrument, with good string-to-resonance coupling also playing an important role. This particular kokle shows outstanding coupling of strings to body-resonances, with all strings evidencing some degree of coupling. In some cases, the strings couple to a superposition of more than one body-resonance.

3:30–3:45 Break

3:45

4pSA4. Modal analysis and computer aided design studies of acoustic guitar structural designs. Craig M. Rashkow, Robert D. Collier (Thayer School of Eng., Dartmouth College, Hanover, NH 03755), and Alan Carruth (Luthier, Newport, NH 03773)

This study seeks to systematically identify modern Computer Aided Design (CAD) techniques to optimize acoustic guitar design. The study uses traditional design as a starting point for analysis of variations in materials and bracing structures of the front sound plate. CAD techniques are used to model the guitar and provide a broadband input that allows analysis of its modal response. Design optimization occurs in creating the most favorable modal response of the instrument while still maintaining structural integrity through effective support of the forces created by the strings' tension in the bridge and neck. The measures chosen to help identify the most favorable response are modal density, magnitude of the response at each mode, and the locations of the largest magnitudes. Increased modal density provides a richer timbre while the amplitude of the response determines the relative intensity of the sound projected. Increased modal response in the upper bout of the instrument can improve the timbre of pitches in the upper frequency range of the instrument. After analyzing these measures for multiple variations and iterations of a selected set of structural parameters, an optimal design is suggested. The results show good agreement between experimental modal measurements and computer-aided design modeling.

4:00

4pSA5. Applying vibration and acoustical radiation from musical instruments in a recording studio setting. Pamela J. Harght (Berklee College of Music, 1140 Boylston St., Boston, MA 02115)

Radiation patterns affect the musical characteristics of any instrument. Applying the radiation characteristics of instruments to microphone technique will enable musicians, audio engineers and students to capture the most-desirable sound, and understand how to position microphones efficiently and appropriately. This paper will use the research done on several instruments' acoustical characteristics and apply it to microphone techniques in a professional-level recording studio at Berklee College of Mu-

sic. Audio examples will be used to demonstrate microphone placement according to the acoustical radiation of musical instruments. [Thank-you to K. Anthony Hoover for his support and encouragement with this paper.]

4:15

4pSA6. Achieving pseudo-degeneracy in handbell modes. John R. Buschert, Sungdo Cha, Daniel A. King, and Daniel B. Horst (Goshen College, 1700 S. Main, Goshen, IN 46526)

Degenerate modes in handbells are split by nonuniformity in the bell. By adding small masses to the bell, the modes can be shifted in position and frequency. Interesting things happen when one attempts to recreate a degeneracy that has been split. Holographic photos of the modes will be shown which can be used to follow the effects of the added mass on individual modes. By moving and varying the mass, one can bring the modes back to a pseudo-degeneracy. Holographic interferometry photos and graphs of the position and frequency of the modes show the stages of this transition.

4:30

4pSA7. Vibration study of Indian Gong hung at one point near the edge. Paresh Shrivage, S. Parmeswaran, and Keith deSa (Acoust. Res. Lab, Dept of Phys., N. Wadia College, Pune, India)

Indian Gong is made up of brass and circular in nature. It is Percussion Instrument. It is used in some religious processions and music concerts in India. It is hung by a metal wire through a hole near the edge of circumference. It is vibrated by a metal (iron) hammer after repeated intervals to get sound of same pitch. This paper relates to vibrational analysis of Indian Gong. A study has been done to check the vibrational properties and the modes of vibration of the Gong. The study is done by spectrum analysis (FFT) and Time-average holography as a vibration analysis tool. The study will help to analyze certain vibrational features of brass. The tonal quality of brass is also compared to other metals, so that it will be useful in making musical instruments. The analysis of hologram may yield some interesting properties of the plate. In conclusion the paper is going to deal with modes of vibration of Indian Gong. The vibrational properties will depend on the point of actuations. The paper describes the mode of vibration and its mathematical equation.

Session 4pSC

Speech Communication: Speech Production and Perception I (Poster Session)

Abeer Alwan, Chair

Dept. of Electrical Engineering, UCLA, 405 Hilgard, Los Angeles, CA 90095

Contributed Papers

All posters will be on display from 1:30 p.m. to 5:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:30 p.m. to 3:15 p.m. and contributors of even-numbered papers will be at their posters from 3:15 p.m. to 5:00 p.m.

4pSC1. Tuning phenomena of melodies and resonance frequencies (formants) in infants pre-speech utterances. Kathleen Wermke (Univ. Wuerzburg, Pre-speech Ctr., Dept. of Orthodontics, Pleicherwall 2, 97070 Wuerzburg, Germany, wermke_k@klinik.uni-wuerzburg.de), Werner Mende (Berlin-Brandenburg Acad. of Sci., Berlin, Germany), Claudia Manfredi, Pierro Bruscoloni (Univ. of Firenze, Italy), and Angelika Stellzig-Eisenhauer (Julius-Maximilians-Univ., Wuerzburg, Germany)

In a former study [Wermke *et al.*, *Med. Eng. Phys.* **24**, 501–514 (2002)], an increasing tuning between laryngeal (melody) and pharyngeal (resonance frequencies) activity was demonstrated during pre-speech development. This tuning was observed unexpectedly early during development in mitigated cries and earliest non-cry utterances and prepares probably articulation in speech-like vocalizations of older infants. A new retrospective study supported this assumption by comparing tuning processes in cries and early non-cry utterances ($N=2500$) in two groups (low versus high word production performance at 18 months) of term-born healthy infants ($N=20$). Additionally, age-matched comparisons were made in 4 cleft-lip-palate-infants treated with a palatal plate. In order to demonstrate the interaction between melody and resonance control we designed a special graphical representation. The tracking function of the resonance frequencies is displayed synchronously to the melody and its harmonics. Resonance frequencies in pre-speech utterances are not yet identical to formant bands associated with speech sounds. The results support the existence of an early active tuning and its relation to later speech. This behavior seems to prepare formant tuning in later speech. Medical applications are seen for infants with disturbances of the vocal tract transfer function, e.g., infants with cleft-lip-palate. [Work supported by DFG.]

4pSC2. The gradient influence of talker sex and perceived sexual orientation on fricative perception. Benjamin Munson, Sarah V. Jefferson, and Elizabeth C. McDonald (Dept. Speech-Lang.-Hearing Sci., Univ. of Minnesota, 115 Shevlin Hall, 164 Pillsbury Dr., SE, Minneapolis, MN 55455, Munso005@umn.edu)

Listeners are more likely to hear a synthetic fricative ambiguous between /s/ and /ʃ/ as /ʃ/ if it is appended to a woman's voice than a man's voice [Strand, *J. Lang. Soc. Psych.* **18**, 86–99 (1999)], suggesting that speech perception is sensitive to social-indexical information. This study examined the influence of two variables on listeners' fricative perception: (a) talker sex, and (b) talkers' perceived sexual orientation (PSO, i.e., the probability that a talker is identified as gay, lesbian, or bisexual (GLB) based on speech alone). Stimuli were created by pairing a synthetic nine-step /s/-to-/ʃ/ series with tokens of /æ/ and /ɪ/ taken from natural productions of *shack* and *ship* by 44 talkers (22 women, 22 men), for a total of 88 different continua. Forty listeners rated the 44 talkers' PSO. A different group of 10 listeners participated in a series of two-alternative *sack-shack* and *sip-ship* identification experiments. As expected, listeners identified more /ʃ/ tokens for women's voices than for men's. GLB-sounding

women elicited significantly fewer /ʃ/ percepts than heterosexual-sounding women. No consistent influence of PSO on fricative identification was noted for men's voices. Regression analyses showed strong relationships between fricative identification and ratings of PSO for women talkers only.

4pSC3. Recognition of English phonemes in noise. José R. Benki (Kresge Hearing Res. Inst., Dept. of Otolaryngol., Univ. of Michigan Med. School, Ann Arbor, MI 48109-1346, benki@umich.edu) and Robert Felty (Univ. of Michigan, Ann Arbor, MI 48109-1285)

Native speakers of American English identified the consonants and vowels of American English CV and VC syllables and the vowels of hVd syllables. The CV and VC syllables, consisting of all of the English consonant phonemes, were selected from recordings described by Shannon *et al.* [*J. Acoust. Soc. Am.* **106**, L71–L74 (1999)], and the hVd syllables were selected from recordings described by Hillenbrand *et al.* [*J. Acoust. Soc. Am.* **97**, 3623–3641 (1995)]. All syllables were presented binaurally in signal-correlated noise. The resulting confusion matrices and feature analyses will be presented along with a position (initial versus final) analysis for the consonant results. The consonant and vowel results will be used to derive empirical measures of similarity of English words. [Work supported by NIH/NIDCD.]

4pSC4. Perturbed palatal shape and North American English /r/ production. Mark K. Tiede (Haskins Labs, 270 Crown St., New Haven, CT 06511, tiede@haskins.yale.edu and M.I.T.-R.L.E., Cambridge, MA), Vincent L. Gracco, Douglas M. Shiller (McGill U., Montreal, QC, Canada), Carol Espy-Wilson (U. of Maryland, College Park, MD), and Suzanne E. Boyce (U. of Cincinnati, Cincinnati, OH)

It is well established that the lowered $F3$ associated with the acoustic percept of American English /r/ can be achieved with different tongue shapes in production. Broadly speaking these shapes may be grouped into "bunched" and "retroflex" varieties. In this work the effects of somatosensory perturbation on /r/ production are examined. Subjects were fitted with a custom palatal prosthesis incorporating a 0.5 cm protrusion along the alveolar ridge, and tongue position during production of /r/ in vocalic contexts was observed using EMA under four conditions: before prosthesis placement; while wearing the prosthesis immediately following placement; still wearing the prosthesis following an unrecorded 20 min adaptation period; and immediately after prosthesis removal. Acoustic effects of the perturbation were minimal, especially after adaptation; production effects were most pronounced in the low vowel context. One subject showed an unperturbed preference for a retroflex configuration, but increased the degree of retroflexion with the palatal prosthesis in place. The remaining subjects preferred unperturbed bunched shapes, but under the

perturbed conditions produced primarily retroflex configurations, which for one subject persisted after prosthesis removal. These results suggest that speaker preference for one shape over another may be determined by palatal morphology. [Work supported by NIH.]

4pSC5. Relational acoustic invariance in the single/geminate stop distinction in Japanese. Yukari Hirata and Jacob Whiton (Dept. of EALL, Colgate Univ., 13 Oak Dr., Hamilton, NY 13346)

This study addressed an issue in the theory of relational acoustic invariance [Pickett *et al.*, *Phonetica* **56**, 135–157 (1999)]. The question was whether an invariant acoustic property exists for distinguishing Japanese single and geminate voiceless stops across different speaking rates. Four native Japanese speakers produced disyllabic words with single and geminate voiceless stops spoken in a carrier sentence at three speaking rates. Durations of sentences, words, stop closures, and vowels preceding the contrasting stops were measured. Ratios of geminate to single stop closures, geminate words to single words, closures to preceding vowels, and closures to words were calculated. The stop closure duration significantly overlapped between the single and geminate categories across rates. However, the ratio of geminate to single closure duration was unaffected by rate. Among the measures examined, the ratio of closure to word duration (0.35 as an optimal boundary) best classified all single and geminate tokens with 95.7–98% accuracy. These results suggest that, in spite of overlap in absolute closure duration between single and geminate stops, there is a relationally invariant measure that divides the two phonemic categories across rates and speakers, supporting the theory of relational acoustic invariance.

4pSC6. Articulatory influences on the categorization of speech sounds. H. Henny Yeung (Dept. of Psych., Univ. of British Columbia, 2136 West Mall, Vancouver, BC, Canada V6T 1Z4, hhyeung@psych.ubc.ca), Bryan W. Gick (Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z4), and Janet F. Werker (Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z4)

This cross-modal study investigates whether production of speech gestures can influence auditory speech perception. Participants made categorization judgments on a continuum between /ba/ and /da/, while making concurrent motor gestures. Ten participants categorized synthesized sounds from a /ba/-/da/ continuum while (a) simply listening, (b) concurrently making a /ba/, /da/, or /ga/ gesture without vocal fold vibration, or (c) making a nonspeech gesture (i.e., tongue protrusion). Preliminary data indicate that making /ba/ and /ga/ gestures, or tongue protrusions, will shift the perceptual boundary away from the /da/ end of the continuum (i.e., towards the /ba/ end of the continuum), compared to a passive listening condition. Making /da/ gestures did not shift the perceptual boundary, compared to this same listening condition. These data offer support for the idea that auditory categorization of speech tokens on a continuum can be influenced by conflicting information in other modalities—specifically articulatory gestures. Further studies will be conducted to examine the influence of (a) other speech gestures on these synthesized sounds, and (b) speech gestures on the perception of naturally-produced speech tokens.

4pSC7. Does intonation have primitive units? Amebu K. Seddoh (Dept. of Commun. Sci. & Disord., Univ. of North Dakota, P.O. Box 8040, Grand Forks, ND 58202)

This study investigated whether listeners can interpret intonation based on partial rather than global fundamental frequency (F_0) or pitch information. Matched Statements and echo questions were recorded digitally at 22 kHz sampling rate using a Computerized Speech Lab (CSL). Each production was stored on the CSL and edited by slicing the wave form into two parts that corresponded to terminal (last 150–200 ms span of the F_0 contour) and preterminal (section of contour preceding the terminal re-

gion) F_0 contour regions. A total of 92 stimuli was thus generated and recorded on a DAT tape. The tape was played to 96 young (20–28 years) and old (42–79 years) adults to determine if each stimulus was part of a statement or a question. For preterminal (PTL) stimuli, both groups successfully identified statements 91% and 88% of the time, respectively, compared to 72% each for questions. The respective scores for the terminal (TL) stimuli were 94% and 91% (statements) versus 91% and 87% (questions). Across groups, statements were easier to identify with both TL ($p=0.002$) and PTL ($p=0.001$) cues compared to questions. These findings suggest that intonation decoding might involve units smaller than global F_0 contour. [Work supported by NIH (NIDCD) grant 5R03DC04955-02.]

4pSC8. Differential effects of speaking rate and phonemic vowel length on formant frequencies of Japanese vowels. Kimiko Tsukada (MARCS Auditory Labs., Univ. of Western Sydney, Penrith South DC NSW 1797 Australia) and Yukari Hirata (Colgate Univ., Hamilton, NY 13346)

This study examined the extent to which variation in speaking rate and phonemic vowel length affect the first two formant frequencies of Japanese vowels. If vowel duration is a factor that determines the degree of formant undershoot [Moon and Lindblom, *J. Acoust. Soc. Am.* **96**, 40–55 (1994)], formants of phonemic long vowels or vowels spoken at slower rates would occupy more peripheral areas in the F_1/F_2 vowel space than the short counterparts. Four male native speakers produced Japanese disyllabic non-words, /mVmV/, /mVVmV/, and /mVmVV/ (V=/i e a o u/, VV=/i: e: a: o: u:/) spoken in a carrier sentence at slow, normal, and fast rates. Effects of vowel length were clear for all five vowels: the long vowels occupied more peripheral areas of the vowel space than the short vowels. However, such a systematic effect was not found with speaking rate changes. ANOVAs performed on each vowel type indicated that, for non-high vowels /e(:) a(:) o(:)/, effects of vowel length and speaking rate interacted. The F_2 of /e:/, for example, did not differ across rates, but the F_2 of /e/ was significantly lower for faster rates. Implications are discussed in terms of Moon & Lindbloms formant undershoot model.

4pSC9. Enhancement of visual perception of speech via tactile input. Diana Gibrael, Bryan Gick, Yoko Ikegami, Kristin Johannsdottir (UBC, Dept. of Linguist., E-270, 1866 Main Mall, Vancouver, BC, Canada V6T 1Z1, gibrael@interchange.ubc.ca), and Jeff Muehlbauer (UBC, Vancouver, BC, Canada V6T 1Z1)

Motor theories of speech perception predict that perceptual information from modalities other than sound enhance speech perception directly by informing the perceiver of the speaker's gestures [A. Lieberman and I. Mattingly, *Cognition* **21**, 1–36 (1985)]. Acoustic theories predict that perceptual information from modalities other than sound will only enhance speech perception if there is a learned mapping between the acoustic speech signal speech and that modality [R. Diehl and K. Kluender, *Ecol. Psych.* **1**, 121–144 (1989)]. As normal subjects are unlikely to have learned a mapping between visual and tactile speech information, this study tests whether and how tactile input enhances visual speech perception. In the control condition, perceivers in noise repeat syllables pronounced by a speaker who they can see clearly. Accuracy is judged on the basis of the repeated syllables. In the experimental condition, subjects additionally have their hand on the speaker's face in the Tadoma position. Results show that subjects are significantly more accurate at perceiving speech when they have both visual and tactile input than when they have visual input alone. In particular, tactile input enhances perceptual accuracy of voice and manner features. [Work supported by NSERC.]

4pSC10. A model of invariant patterns of articulatory movements.

Patrizia Bonaventura (Dept. of Linguist., Univ. of British Columbia, 1866 Main Mall, Buchanan E270, Vancouver, BC, Canada V6T 1Z1, bonaventura.8@osu.edu)

The goal of the present study is to model the “iceberg” portions of the demisyllables, previously extracted from the microbeam articulatory data (Bonaventura, 2003), by curve fitting. The polynomial analysis was designed to include an appropriate weighting window centering around the threshold crossing point, and aimed to provide an estimate of how, in the vicinity of the crossing point, the curve deviates from a straight line: this deviation would be represented by the higher order coefficients of the polynomial. The model was obtained preliminarily on the basis of 100 curves for the lower lip movement for /f/ and /v/ (in initial and final demisyllable for “five”), and from 100 curves for the tongue tip displacement (for /n/ in “nine”). In order to fit the data to the model, a robust least square method (Least Absolute Residuals) has been used, in order to minimize the influence of the outliers, that are present in the read speech data, and cannot be accounted for by “phrase final lengthening effects.” The fit results for the cubic polynomials satisfactorily approximated the “iceberg” curves. The 95% confidence bounds on the fitted coefficients indicated that they were acceptably accurate.

4pSC11. The effects of auditory-visual vowel and consonant training on speechreading performance. Carolyn Richie and Diane Kewley-Port (Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405, crichie@butler.edu)

Recent work examined the effects of a novel approach to speechreading training using vowels, for normal-hearing listeners tested in masking noise [C. Richie and D. Kewley-Port, *J. Acoust. Soc. Am.* **114**, 2337 (2003)]. That study showed significant improvements in sentence-level speechreading for listeners trained on vowels compared to untrained listeners. The present study examined the effects of combining vowel and consonant training on speechreading abilities. Normal-hearing adults were tested in auditory-visual conditions in noise designed to simulate a hearing loss. Using a monosyllable context, one group of listeners received training on consonants, and another group received training on consonants and vowels combined. A control group did not receive training. All listeners performed speechreading pre- and post-tests, on words and sentences. Comparison with the earlier study showed posttest sentence-level speechreading increased by 10 percentage points for participants in the vowel training program, 8 percentage points for participants in the consonant training program, and, unexpectedly, only 2 percentage points for participants in the combined training program. Results from these relatively short training programs suggest that vowels, previously unused in speechreading training, may provide gains in speechreading abilities and play an important role in rehabilitation of hearing-impaired persons. [Work supported by NIHDC02229.]

4pSC12. Turning speech into music in a two-dimensional space by varying the bandwidth and rate of tone pulses placed along formant tracks. C. J. Darwin (Dept. of Psych., Univ. of Sussex, Brighton, BN1 9QG, UK)

A two-dimensional space is introduced which changes sounds from speech-like to music-like by varying the rate and the bandwidth (decay-time) of individual tones placed along formant contours. At one extreme of a very fast pulse rate and a long decay time, the sounds are equivalent to sine-wave speech. With a more moderate tone rate and a short decay time, the sounds become monotone (at the tone-rate) speech. For tone rates around 10/s with long decay times, individual musical tones are heard at the formant frequencies. Intelligibility is poor for such sounds because of the low sampling rate of the formant contours. Increasing the tone rate and decreasing tone decay time increases intelligibility and makes the sounds less music like. Intelligibility data together with speech-like and music-like ratings of sounds from this space will be presented.

4pSC13. Variability of oronasal coupling in children and adults.

Lakshmi Venkatesh and Christopher A. Moore (Dept. of Speech & Hearing Sci., Univ. of Washington, 1417 NE, 42nd St., Seattle, WA 98105)

Oronasal coupling in young children remains poorly understood primarily because of experimental challenges associated with non-invasive techniques for transducing velopharyngeal movement. Productions of the word [bama] were studied in children with normal speech acquisition (NSA) and children with speech delay (SD; 3-to-5 years of age) with respect to adults. Nasal vibration was transduced using a small, light-weight accelerometer affixed to the lateral alar cartilage of the participant's nose. The variability in the nasal acceleration signals obtained for multiple productions of the stimuli for each participant was described using the spatiotemporal index (STI) and functional data analysis (FDA). The spatiotemporal index served as a composite method of variability following linear time normalization of the signals. The FDA technique involving nonlinear time normalization allowed for independent evaluation of amplitude and phase variability. The two groups of children with NSA and SD were more variable than the adults on both measures of STI and amplitude variability, while the amount of phase variability was similar between children and adults. These methods have complementary merits and potential in revealing the properties of development and disruption of oronasal coupling in children and adults. [Work supported by NIH-NIDCD.]

4pSC14. Final devoicing in Russian: Acoustic evidence of incomplete neutralization. Olga Dmitrieva (Dept. of Linguist., Univ. of Kansas, 1541 Lilac Ln., Blake Hall, Rm. 427, Lawrence, KS 66044-3177, olga@ku.edu)

A number of acoustic and perceptual studies conducted on German, Polish, Catalan, and Dutch found evidence of incomplete neutralization of the voicing contrast of obstruents in word-final position. The present study investigates the acoustic correlates of word-final stops and fricatives in Russian. 34 minimal pairs differing in the underlying voicing characteristics of the final segment were incorporated into the body of filler items organized as a stream of associations. This word list was presented to 14 native speakers of Russian. Measurements were obtained for the duration of the vowel preceding the final obstruent, the closure/frication portion of the final obstruent, the burst for the final stops, and duration of the voicing into closure/frication of the final obstruent. Statistical analysis revealed that the effect of underlying voicing, as well as manner of articulation, was significant for all parameters measured. The results strongly suggest that final devoicing in Russian represents a case of incomplete neutralization, at least in the experimental conditions employed.

4pSC15. Haptic-auditory interference from air flow in speech perception. Yoko Ikegami, Diana Gibrael, Bryan Gick, and Kristin Johannsdottir (Dept. of Linguist., The Univ. of British Columbia, E270-1866 Main Mall, Vancouver, BC, Canada V6T 1Z1)

Previous work on haptic interference in auditory perception has shown McGurk-like effects from manual-tactile contact with the face [Fowler and Dekle, *JEP:HPP* **17**, 816–828 (1991)]. The present study investigates whether indirect haptic input affects auditory perception. A novel method was developed in which one experimenter blew puffs of air onto a subject's neck while another produced English plosives, creating possible mismatches between the sensation of aspiration and its acoustic presence. Subjects were blindfolded and wore headphones playing white noise. For 50% of the trial, one experimenter, whose presence was hidden from the subjects, blew puffs of air on subjects' necks lightly enough to be perceived but not noticeable as unnatural. Simultaneously, a second experimenter produced syllables with bilabial plosive onsets (aspirated /p/ or unaspirated /b/) and subjects were asked to repeat what was heard. Sessions were videotaped and three observers rated successful simultaneity of stimuli. Results indicate cross-modal interference. Subjects showed higher accuracy of speech perception when appropriate haptic stimuli accompanied the auditory stimuli. Moreover, when presented with mismatched

tactile and auditory stimuli, subjects demonstrated the fusion of the two modes. Subjects perceived /pa/ when auditory stimulus /ba/ was presented with emulating aspiration. [Work supported by NSERC.]

4pSC16. The (non) categorical perception of place assimilated coronal stops. David W. Gow, Jr. (Cognit./Behavioral Neurology Group, Massachusetts General Hospital, Ste. 340, 175 Cambridge St., Boston, MA 02114, gow@helix.mgh.harvard.edu)

English coronal place assimilation generally produces gradient modification of stop place cues. Acoustic analyses are presented that show that labial assimilation of coronal stops often produces a distinctive pattern of formant movement, hereafter referred to as the coronal step. The coronal step is characterized by an initial pattern of $F1$, $F2$, and $F3$ movement consistent with coronal closure, followed by a second wave of movement producing formant values at offset that are intermediate between those associated with coronal and labial stops. The perceptual consequences of this pattern were examined in categorization and 4I2AFC discrimination tasks using a linear /t/-p/ synthetic VC continuum, as well as a continuum displaying a coronal step but produced by manipulating the same acoustic parameters. The linear continuum produced data consistent with strong categorical perception, while the stepped continuum showed no evidence of categorical perception. These results are discussed in the context of a model of the perception of assimilated speech that relies on the simultaneous activation of competing phonetic categories by assimilated segments. [Work supported by the NIH.]

4pSC17. Position and place effects in Russian word-initial and word-medial stop clusters. Alexei Kochetov (Dept. of Linguist., Simon Fraser Univ., 8888 Univ. Dr., Burnaby, BC, Canada V5A 1S6, alexei_kochetov@sfu.ca) and Louis Goldstein (Haskins Labs./Yale Univ., New Haven, CT 06511-6695)

Studies of inter-gestural timing have shown that (i) word-initial obstruent clusters tend to exhibit less gestural overlap than word-medial or word-boundary clusters, and (ii) the degree of overlap is further affected by the place of articulation of the obstruents (Byrd, 1996; Chitoran, Goldstein, and Byrd, 2002). Both findings have been attributed to perceptual recoverability considerations. This paper presents results of a magnetic articulometer (EMMA) study of Russian word-initial and word-medial stop clusters (e.g., [pt]ashka little bird versus la[pt]a bat). Data collected from 3 native speakers of Russian show that clusters with coronals and dorsals as C1 ([tk], [kt], [kp], [tjm]/[djb]) exhibit less overlap word-initially than word-medially, while the cluster with the labial as C1 ([pt]) does not exhibit the same timing pattern. The findings are interpreted as providing additional support for the role of perceptual recoverability in intergestural timing. First, less overlap in word-initial clusters, compared to word-medial clusters, ensures better place recoverability of C1 (cf. Chitoran *et al.*, 2002). Second, unreleased labials are more perceptually robust than unreleased coronals (Byrd, 1992; Surprenant and Goldstein 1998) and dorsals (Wright, 2001; Kochetov and So, 2005), and thus do not require the same degree of overlap. [Work supported by SSHRC.]

4pSC18. Self-organizing maps for measuring similarity of audio-visual speech percepts. Hans-Heinrich Bothe (Tech. Univ. of Denmark, Ctr. for Appl. Hearing Res., Oerstedts Plads b.352, DK-2800 Lyngby, Denmark, hhb@oersted.dtu.dk)

The goal of this work is to find a way to measure similarity of audio-visual speech percepts. Phoneme-related self-organizing maps (SOM) with a rectangular basis are trained with data material from a (labeled) video film. For the training, a combination of auditory speech features and corresponding visual lip features is used. Phoneme-related receptive fields result on the SOM basis; they are speaker dependent and show individual locations and strain. Overlapping main slopes indicate a high similarity of respective units; distortion or extra peaks originate from the influence of

other units. Dependent on the training data, these other units may also be contextually immediate neighboring units. The poster demonstrates the idea with text material spoken by one individual subject using a set of simple audio-visual features. The data material for the training process consists of 44 labeled sentences in German with a balanced phoneme repertoire. As a result it can be stated that (i) the SOM can be trained to map auditory and visual features in a topology-preserving way and (ii) they show strain due to the influence of other audio-visual units. The SOM can be used to measure similarity amongst audio-visual speech percepts and to measure coarticulatory effects.

4pSC19. Perception of “asymmetrical” German vowels by humans, monkeys and gerbils. Joan M. Sinnott, Leigh Ann Long, and Allen C. Ernst (Psych. Dept., Univ. of South Alabama, Mobile, AL 36688)

Polka and Bohn [J. Acoust. Soc. Am. **100**, 557–592 (1996)] report that human infants perceive certain vowels asymmetrically when tested with a repeating-background category-change operant headturn procedure. Specifically, discrimination is easier when the background is a more central vowel (e.g. /y/) and the target is a more peripheral vowel (e.g. /u/), compared to the opposite condition. We are testing human (adult native English listeners) and monkey discrimination of the German vowel pair /dut-dyt/, using both synthetic and natural stimuli (obtained from Polka and Bohn), and a within-subject design. So far, no significant asymmetries have emerged in the data using a percent correct measure. We will also report preliminary data from gerbils being trained in a between-subject design more comparable to that used with human infants. Plans are to continue testing all species until they reach asymptotic performance levels in order to determine if an RT measure will indicate vowel asymmetries. A video will be shown of humans, monkeys and gerbils working on the experiments.

4pSC20. Effects of reduced orosensory feedback on spectral characteristics of medioalveolar /s/. Juha-Pertti Laaksonen, Stina Ojala, Olli Aaltonen (Dept. of Phonet., Univ. of Turku, FIN-20014 Turku, Finland, juhlaa@utu.fi), Matti Niemi, and Risto-Pekka Happonen (Univ. of Turku, FIN-20520 Turku, Finland)

Effects of reduced orosensory feedback on the production of fricative sounds were studied by measuring spectral characteristics of medioalveolar /s/. Five Finnish male speakers produced sibilant /s/ embedded in 8 different word contexts under normal condition and under reduced condition, in which the tactile information from the tongue was reduced by blocking the lingual nerve on the right side by local anesthesia. Parameters of Long Time Average (LTA) spectrum (i.e., center of gravity, standard deviation, skewness, and kurtosis) were measured by Praat software for every speaker. In comparison between the two conditions, center of gravity, standard deviation, skewness, and kurtosis changed for every speaker. However, the changes were variable and individual. The results of acoustic analysis show that reduced tactile sensation have effects on tongue function producing spectral alterations for sibilant /s/. The inter-individual variation between different talkers suggests that there are no general compensatory mechanisms of speech production, but the mechanisms are highly speaker-dependent.

4pSC21. BMSA, a new Bayesian model selection criterion for assessing audiovisual models of speech perception. J. L. Schwartz (Institut de la Commun. Parle, UMR 5009 CNRS, INPG, Université Stendhal, 46 Av. Flix Viallet, 38031 Grenoble Cedex 1, France)

Audiovisual speech perception has provided matter for many model comparison and assessment studies. Most involved “Root Mean Square Error” (RMSE), a criterion based on the differences between predicted and observed probabilities of response in auditory, visual and audiovisual categorization experiments. RMSE is related to the best fit of a model considering data. However, some models are known for their ability to

adapt to almost any data set. This is always associated to a large fit instability: very small variations of the model parameters lead to dramatic modifications of the predicted data. A new criterion is derived from the Bayesian Theory of model selection. Considering that best fit estimation is not error-free, model comparison in this theory involves the total likelihood of the model knowing the data, integrating on the whole parameter space. This results in penalizing a too flexible model, for which local likelihood is high, but global likelihood is low. An approximation of this criterion, called BMSA, is presented. This criterion, easy to compute, results in combining fit and stability. An application of the BMSA criterion in an audiovisual speech perception experiment is proposed, enabling to show for the first time that the fusion process is subject-dependent.

4pSC22. Age differences in detecting gaps in speech and non-speech.

M. Kathleen Pichora-Fuller, Bruce A. Schneider (Dept. of Psych., Univ. of Toronto, 3359 Mississauga Rd., Mississauga, ON, Canada L5L 1C6, kpfuller@utm.utoronto.ca), Nancy Benson (The Hospital for Sick Children, Toronto, ON, Canada M5G 1X8), Stanley Hamstra (Univ. of Toronto, Toronto, ON, Canada M5G 1L5), and Edward Storzer (McNeill Audiol., Victoria, BC, Canada V8R 1G1)

Ability to detect gaps in speech and non-speech stimuli was measured in children, young adults, and older adults with good audiograms. The markers varied in duration (40 vs 250 msec) and in spectral symmetry. In spectrally symmetrical conditions, the leading and lagging markers were the same: the vowel [u] in speech conditions and a 500-Hz tone in non-speech conditions. In asymmetrical speech conditions, the lagging marker was the same as in the symmetrical conditions, but the leading marker was the consonant [s] in the speech conditions and a broadband noise (1 to 6 kHz) in the non-speech conditions. For all groups, gap detection thresholds in spectrally symmetrical markers were far smaller than in spectrally asymmetrical markers. Thresholds were significantly smaller in young adults than in either children or older adults. Gaps between spectrally asymmetrical speech markers were detected better than gaps between analogous non-speech stimuli. It is argued that phonological knowledge compensates for auditory processing difficulties. [Research funded by the International Dyslexia Association, the Natural Sciences and Engineering Research Council of Canada, and the Canadian Institutes of Health Research.]

4pSC23. Voice onset time in Mandarin esophageal speech.

Manwa Ng (Long Island Univ., 720 Northern Blvd., Brookville, NY 11548-1300, manwa.ng@liu.edu) and Hanjun Liu (Xi'an Jiaotong Univ., Xi'an, 710049, P. R. China)

As an important perceptual cue for voicing and aspiration of stops, voice onset time (VOT) is mainly determined by the aerodynamic interaction between the intraoral and subglottal regions. However, since the PE segment serves as a new vibratory source in esophageal phonation, aerodynamic events are very different from laryngeal phonation. VOT associated with esophageal speech of English has been reported previously. However, few studies have reported VOT characteristics of esophageal speech of tone languages. The present study will investigate the possible VOT difference between esophageal and normal laryngeal speakers of Mandarin Chinese. Seven superior esophageal speakers and 7 normal laryngeal speakers will participate in the present investigation. They will be native male speakers of Mandarin Chinese. The participants will produce the syllable /ta/ embedded in a carrier phrase at a comfortable loudness level for three times. VOT values will be measured from a time domain waveform. With reference to a wide-band (300 Hz) spectrogram, VOT will be defined as the release of stop and the first identifiable period of the

vocalic portion. An independent sample t-test will be used to determine if VOT values of esophageal and normal laryngeal speakers were significantly different.

4pSC24. A role for tracheal resonances in speech perception.

Asaf Bachrach (Dept. of Linguist. and Philosophy, MIT, Cambridge, MA 02139), Steven Lulich (MIT, Cambridge, MA 02139, lulich@mit.edu), and Nicolas Malyska (MIT, Cambridge, MA 02139)

Acoustic coupling between the vocal tract and the trachea results in the introduction of pole-zero pairs corresponding to resonances of the uncoupled trachea. If the second formant (F_2) passes through the second tracheal resonance (T_2) a discontinuity in amplitude occurs. This work explores the hypothesis that the F_2-T_2 discontinuity affects how listeners perceive the distinctive feature [back] in transitions from a front vowel (high F_2) to a labial stop (low F_2). We synthesized two versions of an utterance ("apter") with an F_2-T_2 discontinuity at different locations in the initial VC transition. Subjects heard portions of the utterance with and without the discontinuity, and were asked to identify the utterance. Results show that the presence of the F_2-T_2 discontinuity facilitated the perception of frontness in the vowel. Discontinuities of the F_2-T_2 sort are proposed to play a role in shaping vowel inventories in the world's languages [K. N. Stevens, J. Phonetics 17, 3-46 (1989)]. Our results support a model of lexical access in which acoustic discontinuities subserve phonological feature identification.

4pSC25. Targetless schwa revisited.

Edward Flemming (Dept. of Linguist. and Philosophy, M.I.T., 77 Massachusetts Ave., 32-D808, Cambridge, MA 02139, flemming@mit.edu)

It has been suggested that schwa vowels in English lack vowel quality targets, that is, they are realized as interpolations between the preceding and following segmental contexts or are the result of separating two consonantal constriction gestures, without an inherent vocalic gesture. Tests of this hypothesis have produced mixed results. The present study revisits the question of schwa targets based on acoustic analysis of schwa vowels in a wider range of segmental and morphological contexts. The main conclusions are: (i) It is important to distinguish stem-final reduced vowels (as in "pandA") from other schwa vowels (e.g., "About," "bEgin"). The former clearly have vowel quality targets they are generally mid central vowels whereas the latter are much more contextually variable. (ii) The variable schwa vowels have a target, but this target is not a particular vowel quality, it is to indicate the presence of a vowel. This is achieved through a variety of means, including duration and realization of a local amplitude peak, but is consistent with substantial variation in vowel quality, particularly F_2 .

4pSC26. The role of attention in infant phonetic perception.

Monika Molnar, Linda Polka, and Susan Rvachew (McGill Univ., 1266 Pine Ave. W., Montreal, QC, H3G 1A8 Canada monika.molnar@mcgill.ca)

Attention is an important factor underlying phonetic perception that is not well understood. In this study we examined the role of auditory attention in infant phonetic perception using a distraction masker paradigm. We tested infant discrimination of /bu/ vs /gu/ with a habituation procedure and three natural productions of each syllable. For the quiet condition each token was copied into a separate sound file. For the distractor condition, a high frequency noise was added to each sound file so that it gated on and off with the onset and offset of the syllable. The distractor noise was a recording of bird and cricket songs whose frequencies did NOT overlap with the test syllables. Thus, the noise did not change the audibility of the syllable, but it could distract infants if they do not focus their attention well. Infants (6- to 8-month-olds) were tested in each condition. Infants tested in quiet performed significantly better than infants tested in the distractor condition; discrimination scores showed little overlap between

the two groups. These findings indicate that in young infants, attention to subtle phonetic differences is easily disrupted. The implications for developmental models of speech perception will be discussed.

4pSC27. The development of laryngeal coarticulation: Comparison of women, 5-year-olds, and 10-year-olds. Laura L. Koenig (Haskins Labs & Long Island Univ., Brooklyn)

Past work has indicated that adult speakers show effects of laryngeal coarticulation in voice source measures taken in vowels flanking voiceless consonants. In a recent pilot study, we presented time-varying voice source measures from 5-year-old girls producing VCV sequences with varying consonants. The results suggested that, on average, the children produced laryngeal coarticulation over a duration at least as long as adults, but extensive token-to-token variability complicated interpretation of the results. The present analysis extends that work by comparing normal women with normally-developing 5-year-old and prepubertal 10-year-old children. Speakers were recorded producing multiple repetitions of intervocalic /b p h/ in simple carrier phrases such as “Poppa Hopper” while oral airflow signals were collected using a Rothenberg mask. After software-inverse filtering, pulse-by-pulse measures of DC airflow, open quotient, and f_0 are made from the transition out of the consonant into the following vowel. Results will be interpreted in terms of age effects and, for the children, possible gender differences. These data have implications for our understanding of laryngeal motor control in children and, more generally, for theories of the development of coarticulation. [Work supported by NIH.]

4pSC28. Tonal and durational variations as phonetic coding for syllable grouping. Yi Xu (Univ. College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, UK) and Maolin Wang (College of Chinese Language and Culture, Jinan Univ., Guangzhou, 510610, China)

While syllables in connected speech are generally believed to be prosodically divided into groups, how such grouping is done phonetically is not fully understood. This study explores the possibility that syllable grouping is partly realized through adjusting articulatory strength and duration. We compared the degrees of tonal undershoot in Mandarin as they are related to syllable position and number of syllables in words or phrases. The sequences consisted of 1–4 syllables with R or F tone and were produced by eight speakers. The all-R and all-F sequences impose great pressure on tone production and hence would best reveal the effects of strength and duration. Results show that as the number of syllables in each sequence increased, both syllable duration and size of F_0 excursion decreased. Meanwhile, excursion size varied with syllable location in the sequence. But in each case it was the first and last syllables that had the largest excursions, and the excursion variations cannot be fully accounted for by duration. There thus appear to be both an isochrony effect and an “edge marking” effect. Taking into consideration known prosodic effects on segmental phonemes, syllable grouping as a communicative function seems to involve a rather complex encoding scheme.

4pSC29. Is speech lazy or just efficient? A control-theoretic analysis. Luis Rodrigues (Dept. of Mech. and Industrial Eng., Concordia Univ., 2160B Bishop St., B-304 Montreal, QC, Canada H3G, luisrod@me.concordia.ca) and John Kroeker (Eliza Corp., Beverly, MA 01915)

This paper introduces a control-theoretic model that allows us to address the energy dynamics of the vocal tract system. The model can be generated directly from articulatory data. The model allows computation of the energy in a state transfer, from an initial to a final articulator configuration. This method can help determine the degree of physical feasibility of various proposed articulatory trajectories. The basic assumption is that the set of articulators evolve through configurations that minimize the energy spent by the system to produce an utterance. Minimum control energy gives a measure of how hard it is to reach a target point for the

different articulators. Simulation results are presented corresponding to the computation of the minimum energy from the MOCHA database. The linear model is shown to be adequate for short, well-labeled segments. The results show the intriguing fact that minimum control energy seems to have an oscillatory (swinging) nature for the production of speech. Physical features such as time constants and natural frequencies of the articulators are derived. A control-theoretic model of the dynamics of the mechanical articulators of speech production could be a fundamental tool to understand the mechanism of speech production.

4pSC30. Effects of simultaneously presented pitch- and loudness-shifted voice auditory feedback on voice fundamental frequency.

Charles R. Larson, Jean Sun, and Hideki Takaso (Northwestern Univ., 2240 Campus Dr., Evanston, IL 60208)

Recent studies have demonstrated that subjects respond to pitch- or loudness-shifted voice feedback with compensatory changes in voice fundamental frequency (F_0) or intensity. In the present study, subjects sustaining /u/ vowel sounds were presented with either pitch-shifted auditory feedback (50 or 100 cents) or simultaneous pitch-shifted and loudness-shifted voice feedback (3 or 6 dB SPL) of 200 ms duration. During simultaneous stimulus presentations, the stimulus directions were opposite in direction (e.g., increasing pitch and decreasing loudness feedback). Subjects ($N=24$, normal, age 18–24) responded to pitch-shifted feedback with compensatory adjustments in F_0 . Subjects responded to simultaneous stimuli with independent compensatory adjustments in voice F_0 and intensity. Latencies of F_0 responses to simultaneous stimuli ($m=211$ ms) were longer than to pitch-shifted stimuli alone ($m=145$ ms, $F=26$, $df=2306$; $p<.0001$). Results suggest auditory feedback mechanisms for stabilizing the direction of voice F_0 against perturbed pitch feedback are independent of mechanisms for stabilizing loudness feedback, but the simultaneous presence of pitch and loudness perturbations slows down the corrective mechanism for pitch-shifted feedback.

4pSC31. Phoneme clustering based on segmental lip configurations in naturally spoken sentences.

Jianxia Xue, Abeer Alwan (Dept. Elec. Eng., Univ. of California, Los Angeles, 405 Hilgard Ave., Los Angeles, CA 90095, jxue@ee.ucla.edu), Jintao Jiang, and Lynne E. Bernstein (House Ear Inst., Los Angeles, CA 90057)

It has been shown that face (lips, cheeks, and chin) information can account to a large extent for visual speech perception in isolated syllables and words. Visual speech synthesis has used small reduced sets of phonemes (“visemes”), under the theory that perceivers are limited in their ability to extract visual speech information. In this study, lip configurations from a manually segmented sentence database [L. Bernstein *et al.*, J. Acoust. Soc. Am. **107**, 2887 (2000)] were analyzed to provide phoneme clusters that are algorithmically distinguishable using mouth vertical/horizontal opening and lip protrusion from the middle position of each segment. The lip feature sample spaces for each phoneme were represented by Gaussian mixture models. Maximum posterior probability classification results were computed for each phoneme. Confusion matrices were generated from the classification results, and a set of confusions with 74% or higher within-group classification correct was judged to be a cluster. Preliminary results from 191 sentences by a single talker generated the following clusters: {/p, b, m/(77%), /f, v/(74%), /w, r/(80%), /t, d, s, z, D, k, n/(88%)}. We will present results analyzing the entire English phoneme set across different talkers and compare the results with visual perceptual clusters. [Work supported in part by the NSF.]

4pSC32. The effects of signal-to-noise ratio on auditory-visual integration: Integration and encoding are not independent. Mitchell S. Sommers (Dept. of Psych., Washington Univ., Campus Box 1125, St. Louis, MO 63130, Msommers@wustl.edu), Brent Spehar, and Nancy Tye-Murray (Washington Univ. School of Medicine, St. Louis, MO 63130)

Most current models of auditory-visual speech perception propose a two-stage process in which unimodal information is extracted independently from each sensory modality and is then combined in a separate integration stage. A central assumption of these models is that integration is a distinct perceptual ability that is separate from the ability to encode unimodal speech information. The purpose of the present study was to evaluate this assumption by measuring integration of the same speech materials across three different signal-to-noise ratios. Twelve participants were presented with 42 repetitions of 13 consonants presented in an /iCi/ environment at 3 different signal-to-noise ratios. Integration was assessed using an optimum processor model [L. Braida, *Q. J. Exp. Psych.* **43A**, 647–677 (1991)] and a new measure termed integration efficiency that is based on a simple probability metric. In contrast to predictions made by current models of auditory-visual speech perception, significant differences were observed for both measures of integration as a function of signal-to-noise ratios. These findings argue against strictly serial models of auditory-visual speech perception and instead support a more interactive architecture in which unimodal encoding interacts with integration abilities to determine overall benefits for bimodal speech perception. [Work supported by NIA.]

4pSC33. The University of South Florida audiovisual phoneme database, v 1.0. Stefan A. Frisch, Sarah Hardin, Dee Adams Nikjeh, and Adrienne M. Stearns (Dept. of Commun. Sci. and Disord., Univ. of South Florida, Tampa, FL 33620, frisch@cas.usf.edu)

An audiovisual database of English speech sounds has been developed for teaching purposes. This database contains a variety of Standard English speech sounds produced in clear speech in isolated words. Phonemes are produced in word initial, word medial, and word final position, unless not allowed by English phonotactics. There is one example of each word spoken by a male talker of Standard American English from the Midwest. The database consists of individual word files that contain simultaneous audio recording, video of the lips, flexible endoscopic image of the pharynx and larynx, and ultrasound video of the tongue in the mid-sagittal plane. The files in the database are uncompressed video avi format and are suitable for examination in the Wavesurfer freeware program (Sjolander and Beskow, KTH Stockholm). This database is intended as a multimedia reference for students in phonetics or speech science. Its coverage is relatively complete, though there are some gaps due to technical difficulties with the recording procedure. A demonstration of selected recordings from the database and suggestions for their use will be presented. Plans for further development of the database will also be discussed. [Work supported by a USF Innovative Teaching Grant.]

4pSC34. Generalization of the imitation effect within a natural class. Kuniko Nielsen (Dept. of Linguist., UCLA, 3125 Campbell Hall, Los Angeles, CA 90095)

A natural class refers to a set of segments which share acoustic or articulatory features. Support for this notion has traditionally been provided by phonological alternations and phonotactic constraints, and recently by experiments (e.g., Goldrick, 2004). This study investigates the psychological reality of natural class using the imitation paradigm (Goldinger, 1998) in which subjects speech is compared before and after they are exposed to target speech (= study phase). Although this paradigm has shown that subjects shift their production in the direction of the target, these results do not reveal the size of the linguistic unit(s) influenced by the effect. That is, when a subject shifts production of a particular sound in a word, it is uncertain whether the subject is picking up on the word, the segment, or the feature. In this study, to investigate whether phonetic

imitation is generalized across members of a natural class, the study-phase word list includes words with initial /p/ and /t/ (with extended VOT), while the pre- and post-study production list includes (1) the modeled words, replicating Shockley *et al.* (2004), (2) the modeled segments /p/ and /t/ in new words, and (3) the modeled feature [+spread glottis] (aspiration) in a new segment /k/.

4pSC35. Functional MRI reveals two distinct cerebral networks subserving speech motor control. Axel Riecker (Dept. of Psychiatry III, Univ. of Ulm, Leimgrubenweg 12-14, D-89075 Ulm, Germany), Klaus Mathiak (Univ. of Aachen, Pauwelsstr. 30, D-52074 Aachen, Germany), Wolfgang Grodd (Univ. of Tuebingen, D-72076 Tuebingen, Germany), Ingo Hertrich, and Hermann Ackermann (Univ. of Tuebingen, D-72076 Tuebingen, Germany, hermann.ackermann@uni-tuebingen.de)

To further delineate the neural basis of speech motor control, functional magnetic resonance imaging (fMRI) was performed during syllable repetitions synchronized to click trains (8 subjects, 2–6 Hz; vs. passive listening task). (a) Bilateral hemodynamic responses emerged within mesiofrontal and sensorimotor cortex, putamen / pallidum, thalamus and cerebellum (two activation spots at either side). In contrast, dorsolateral premotor cortex and anterior insula showed left-sided activation. (b) Calculation of rate / response functions revealed a negative linear relationship between repetition frequency and hemodynamic activation within the striatum whereas both cerebellar hemispheres exhibited a step-wise response increase at about 3 Hz. (c) Analysis of the temporal dynamics of hemodynamic activation revealed these cortical and subcortical brain regions to be organized into two separate networks (medial and dorsolateral premotor cortex, anterior insula, superior cerebellum versus sensorimotor cortex, basal ganglia, inferior cerebellum). These data provide evidence for two levels of speech motor control bound, most presumably, to motor preparation and executions processes. Furthermore, these findings help to explain clinical observations such as an unimpaired or even accelerated speaking rate in Parkinsons disease and slowed speech tempo which does not fall below a rate of about 3 Hz in cerebellar disorders.

4pSC36. A phonetic study of guttural laryngeals with data from Semitic. Kimary Shahin (Effat College of Al-Faisal U., P.O. Box 34689, Jeddah 21478, Saudi Arabia, kshahin@effatcollege.edu.sa)

Phonetic data from Hebrew and Arabic were examined to determine if laryngeal consonants in those languages have tongue root articulation. This was done in search of an articulatory basis for the phonological patterning of Semitic laryngeals with tongue root consonants like pharyngeals and uvulars. This issue has been problematic for phonological theory because a theoretically comfortable phonetic basis for that patterning has not yet been identified [J. J. McCarthy, *Pap. Lab. Phonol.* **III**, 191–234 (1994); F. Nolan, *ibid.* **IV**, 361–367 (1995)]. The present study is the first to address the problem with clear natural-language data. Acoustic and articulatory data are presented (digital audio and video, waveform, spectrogram and video picture) from two laryngoscopic films: Prof. A. Laufer's film of Hebrew and Arabic speech recorded at Haskins Labs in the early 1980s, and a new film of Arabic speech recorded at Speech Technology Research. The data show no tongue root articulation for the laryngeals. The implications for phonology are discussed. Note: Thanks to Prof. Laufer and Haskins Labs for permission to use the Haskins film, and to Prof. John Esling for collaboration in recording the new Arabic data. [Work supported by a SSHRCC grant to J. Esling.]

4pSC37. Quality of American English back vowels before /r/. Michael J. Clark and James M. Hillenbrand (Speech and Hearing Ctr., MS5355, Western Michigan Univ., Kalamazoo, MI 49008)

The vowels /o/ and /ɔ/ are not contrastive before /r/ in most American English dialects, and the phonetics literature is equivocal about the phonetic quality of the nucleus in words such as *board*. Most works use [ɔ] to

represent the first part of the nucleus. In this study acoustic measurements, listening tests, and discriminant analyses were used to determine the phonetic quality of such vowels. Fourteen women recorded monosyllables containing /o,ɔ,a/, with initial /b,g,h/ and final /d/ or /z/ (e.g., *bode*, *Baud*, *bod*). Additionally a set with rhotic diphthongs (*board*, *barred*, *gored*, *guard*, *hoard*, *hard*) was recorded. The central back vowel in rhotic diphthong syllables (*board*, *gored*, *hoard*) showed formant values very similar to those for /o/ and very unlike those for /ɔ/. The low back rhotic diphthong (i.e., *barred*, *guard*, *hard*) showed a range of values from /a/ to /ɔ/. Listening tests using brief excerpts from vowel onsets supported the greater similarity of *board* to *bode* than to *Baud* and the intermediate nature of *barred* between *bod* and *Baud*. Discriminant classification of the rhotic diphthong formant measures supported the same conclusions. [Work supported by NIH.]

4pSC38. Articulatory and acoustic characteristics of English /l/ in children's speech production. Sunyoung Oh and Bryan Gick (Dept. of Linguist., Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z1, oh.sun_young@courrier.uqam.ca)

English /l/ is one of the later-developing sounds in language acquisition. It has syllable-based allophones (e.g., light l in initial, dark l in final) and comprises multiple articulatory gestures (e.g., tongue tip fronting and raising, tongue dorsum backing, lateral dipping) with different coordination in timing and magnitude in syllable position (e.g., synchronicity in initial, tongue tip delay in final with greater reduction in magnitude). Using ultrasound, this study examined how articulatory characteristics of /l/ are presented in children's speech production. Replicating previous articulatory studies in adult speech production, /l/ words in isolation and combination were collected from eight monolingual children aged 3;11 to 5;9. Although some children produced more similar to adult /l/, children produced /l/ using different articulatory properties. /l/ was produced with less

detail and more variations than in the adults' speech as movements of the tongue were simplified or modified. While articulatory performance varied across subjects as well as syllable positions, acoustic analysis of formants ($F1$, $F2$, $F3$) of sample tokens showed similar patterns for all subjects. This study suggests that the tendency toward late acquisition of /l/ is due directly to the articulatory as well as motor complexity.

4pSC39. The effects of tongue shape categories on tongue segmentation in English. Melissa A. Epstein and Maureen Stone (Biomed. Sci., Univ. of Maryland Dental School, Rm. 5A12, 666 W. Baltimore St., Baltimore, MD 21201)

In our recent work, we have proposed that the tongue moves by compressing and expanding local functional segments. For any single gesture, functional segments may move in similar or opposite directions to compress and expand the tongue locally. High correlations between segments suggest biomechanical constraints. Low correlations suggest independent control of these segments. Our previous studies on English have shown a front-back division of the tongue, where adjacent segments correlate positively with each other (moving in the same direction) and distal segments correlate negatively with each other (moving in opposite directions). Individual segments aggregate with adjacent or distal segments. Furthermore, the phonemic content of the dataset influences these aggregations and the location of the front-back division of the tongue. This study will more deeply explore these phonemic effects. In particular, we will examine the effects of the transitions between consonants and vowels of the four basic tongue shape categories of English [front raising (e.g., /n,i/), back raising (e.g., /ng,o/), complete channel (e.g., /s,ae/) and two point displacement (e.g., /l/)] on tongue segmentation and the location of the pivot point for the front-back division. [Work supported in part by NIDCD/NIH Grant RO1-DC01758 and by NIH Grant T32-DE07309.]

THURSDAY AFTERNOON, 19 MAY 2005

GEORGIA B, 2:00 TO 3:45 P.M.

Session 4pUWa

Underwater Acoustics: Propagation: Modeling and Experimental Results II

Kevin B. Smith, Chair

Dept. of Physics, Naval Postgraduate School, Monterey, CA 93943

Contributed Papers

2:00

4pUWa1. Multipath cancellation with a two-channel array. John E. Piper (NSWC-PC, Code R21, 110 Vernon Ave., Panama City, FL 32407, john.e.piper@navy.mil)

A test pool experiment with direct and multipath signals incident on a small vertical array was conducted. To separate these signals a maximum likelihood method approach was used. This approach is based on exploiting the orthogonal nature of the signals in the maximum likelihood parameter space, which leads to a theoretically complete decoupling of the desired signal from the multipath interference. This approach has no analog in conventional signal processing. Results of the test pool experiment using this maximum likelihood method cancellation and conventional beamforming methods are presented.

2:15

4pUWa2. Torpedo detection using multi-path signals and fast orthogonal search techniques. Jeff Collins, Donald McGaughey (Dept. of Phys., Royal Military College of Canada, P.O. Box 17000, Stn Forces Kingston, ON, Canada K7K 7B4, jeff.collins@rmc.ca), Jim Theriault, and Sean Pecknold (Defence Res. and Development Canada (Atlantic), Dartmouth, NS, Canada B2Y 3Z7)

Detecting a high speed torpedo by means of a passive acoustic detector is very challenging for most acoustic operators. Coupled with a very noisy environment, multiple sources in a multi-path scenario and varying environmental factors, a time-constrained assessment will prove difficult. In addition, a passive sensor cannot estimate the range of a torpedo approaching it at a constant bearing. The passive acoustic sensor will receive a direct path signal from the torpedo as well as a signal that has reflected off

4p THU. PM

the surface. Due to the different angle of arrival, the direct-path and surface-reflected signals have different Doppler shifts. The Torpedo Detection Algorithm (TDA) employs the fast orthogonal search (FOS) algorithm for high-resolution spectral analysis to detect the closely spaced direct-path and surface-reflection signals. When a direct-path and surface-reflection are found, an automatic alert of a torpedo detection is initiated. In simulation, a torpedo is detected 20 times out of 20 as it travels from 5000 to 500 m from the receiver. Simple trigonometric expressions are used to estimate the torpedos range given the two frequencies estimated by FOS and *a priori* information about the torpedo speed and depth.

2:30

4pUWa3. A comparison of mine counter measure performance models. Andrew Holden (Dstl. Winfrith, Dorchester, DT2 8WX, UK, apholden@dstl.gov.uk)

Mine counter measure (MCM) sonar systems perform the task of detection and classification of marine mines that are typically laid in shallow water environments. Currently, there are several MCM performance models in use that can predict the performance of MCM sonars. This paper gives a brief description of some these models and gives a comparison of their performance predictions for several shallow water scenarios. All the models examined are self contained packages that can model the entire problem—the sonar, the environment, and the target. They are considered to be energy models in that only the intensity of sound received from various parts of the environment is modeled while phase calculations are ignored. The results show that the models can give good agreements with each other for some scenarios. In some other scenarios the agreement is not so good and reasons are given to show why this happens.

2:45

4pUWa4. Tank experiments and model comparisons of shallow water acoustics over an elastic bottom. Jon M. Collis, William L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180, collisj@rpi.edu), Michael D. Collins, Erik C. Porse, Harry J. Simpson, and Raymond J. Soukup (Naval Res. Lab., Washington, DC 20375)

A series of tank experiments are being conducted in order to obtain high quality data for acoustic propagation in environments with sloping elastic bottoms. Such problems can now be solved accurately with the parabolic equation method, which is being used to model the experiment. This paper will present results of the initial experiments and discuss plans for upcoming experiments, which will include propagation onto land. The initial experiments involved a broadband source over a block of PVC that was suspended in deionized water. Time series were collected at 100 to 300 kHz on horizontal and vertical arrays for two source positions. [Work supported by the Office of Naval Research.]

3:00

4pUWa5. Adjoint-based control of nonlocal boundary conditions for Claerbout's wide-angle parabolic approximation. Matthias Meyer and Jean-Pierre Hermand (Dept. of Optics and Acoust., Univ. Libre de Bruxelles, av. F-D. Roosevelt 50 - CP 194/05, B-1050 Brussels, Belgium)

This paper applies the concept of optimal boundary control for solving inverse problems in shallow water acoustics. A continuous analytic adjoint model is derived for a wide-angle parabolic equation (WAPE) using a generalized nonlocal impedance condition at the water-bottom interface.

While the potential of adjoint methodology has been demonstrated for ocean acoustic tomography, this approach combines the advantages of exact transparent boundary conditions for the WAPE with the concept of adjoint-based optimal control. In contrast to meta-heuristic approaches the inversion procedure itself is directly controlled by the waveguide physics and, in a numerical implementation based on conjugate gradient optimization, much fewer iterations are required for assessment of environments that are supported by the underlying subbottom model. Furthermore, since regularization is important to enhance performance of full-field acoustic inversion, special attention is devoted to applying penalization methods to the adjoint formalism. Regularization incorporates additional information about the desired solution to stabilize ill-posed problems and identify useful solutions, a feature that is of particular interest for inversion of field data sampled on a vertical array in the presence of measurement noise and modeling uncertainty. Results show that the acoustic fields and the bottom properties embedded in the control parameters are efficiently retrieved.

3:15

4pUWa6. Line-integral prediction for horizontal coherence in deep-water propagation. Michael Vera (Univ. of Southern Mississippi, 730 E. Beach Blvd., Long Beach, MS 39560, michael.vera@usm.edu)

Some of the characteristics of an acoustic signal propagating to basin-scale ranges in the ocean can be estimated using line integrals along deterministic ray paths. These line-integral approximations involve the statistics of the internal-wave field. The success or failure of the integral expressions can be analyzed by comparison to parabolic-equation simulations through multiple realizations of the stochastic internal-wave field. One acoustic characteristic of interest is the length scale of acoustic horizontal (cross-range) coherence. Recent work comparing integral predictions of horizontal coherence length to values from simulations of the North Pacific Acoustic Laboratory experiment yielded mixed results. Additional comparisons will be discussed for acoustic propagation paths that are not impacted by range-dependence in the bathymetry or background sound-speed profile. These simulations will employ a single sound-speed profile and a deep, range-independent bottom. The comparisons can yield some insight into the accuracy of the integral estimate; they will not be influenced by bathymetric interaction or range dependence in the background sound speed.

3:30

4pUWa7. Decomposition method in constructing simulation models of parametric location for statistically irregular mediums. Irene Starchenko (TSURE, 347928, Taganrog, GSP-17a, Nekrasovskiy, 44, Russia, star@tsure.ru)

In using parametric arrays for the purposes of distant sounding in a water medium it is necessary to take into account the probable characteristics of acoustic signals. In this case the modeling of processes is especially important, because experiments in natural conditions are not always possible. In the case of parametric location the medium plays the very important function of formation of the parametric array. Results will be discussed.

Session 4pUWb

Underwater Acoustics: Underwater Noise Studies

David E. Hannay, Chair

JASCO Research Ltd., 2101-4464 Markham St., Victoria, BC V8Z 7X8, Canada

Contributed Papers

2:00

4pUWb1. Range estimation of broadband noise sources in an ocean waveguide using the array invariant. Sunwoong Lee and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139, makris@mit.edu)

A method is developed for range estimation of broadband noise sources in a horizontally stratified ocean waveguide without *a priori* knowledge of the environment. It has previously been shown that the range of a transient source can be estimated using the “array invariant” method by analyzing instantaneous beam-time intensity data [Lee and Makris, *J. Acoust. Soc. Am.* **116**, 2646 (2004)]. This method is now extended to range estimation of continuous, broadband noise sources in an ocean waveguide. It is shown that the cross-correlation of instantaneous beam-time intensity asymptotically reaches the array invariant. The range of the source can then be determined without knowledge of the environmental parameters. This method is applied to localize multiple uncorrelated noise sources in a horizontally stratified ocean waveguide without ambiguity.

2:15

4pUWb2. Effect of shallow water internal waves on the ambient noise notch. Daniel Rouseff, Dajun Tang, and Frank S. Henryey (Appl. Phys. Lab., College of Ocean and Fishery Sci., Univ. of Washington, Seattle, WA 98105, rouseff@apl.washington.edu)

Consider a vertical receiving array in shallow water that is recording ambient noise. If the recorded noise is beamformed, the resulting beam pattern often exhibits a notch at broadside to the array. Modeling of range-independent environments suggests that this ambient noise notch should become more pronounced as the sound speed gradient is increased. In the present work, the effect of range dependence on the ambient noise notch in the 1–5 kHz band is studied. Range dependence is introduced into the model in the form of random shallow water internal waves. Coupling between the propagating acoustic modes is calculated using the Dozier-Tappert formulation [*J. Acoust. Soc. Am.* **63**, 353–365 (1978)] extended to consider bottom loss. The model is internally consistent as the buoyancy profile determining the internal waves is dictated by the sound speed gradient. Experimental results from the 2001 East China Sea Experiment are also reported. [Work supported by ONR.]

2:30

4pUWb3. Using vertical directionality to study the ambient noise fluctuations due to internal waves in South China Sea. Hsiang-Chih Chan (Natl. Taiwan Univ., No. 1, Sec. 4, Roosevelt Rd., Taipei, 106, Taiwan, d91525003@ntu.edu.tw), Ruey-Chang Wei, Wen-Zheng Hu (Natl. Sun Yat-sen Univ., Kaohsiang, 804, Taiwan), and Chi-Fang Chen (Natl. Taiwan Univ., Taipei, 106, Taiwan)

A range-dependent acoustic propagation model is used to simulate the complex ambient noise field, which offers the key to the understanding of noise varying during the ASIAEX 2001 SCS experiment. The surface noise sources that are generated by wind/waves and ships are distributed

horizontally in the simulated model. Observed ambient noise fluctuations in South China Sea are greatly affected by prevailing internal waves in this area, the vertical directionality changes on vertical line array output were believed to be results of interaction between sound and ocean environment. The measured twin-peak beam patterns are not consistent with previous studies on flat or upslope bathymetry. Therefore, in this study the discrepancy is explained by modeling and correlation between the noise distributions and internal wave position.

2:45

4pUWb4. Acoustic spreading and boat noise within shallow Florida waterways. Richard Phillips (Dept. of Mech. and Aerosp. Eng., Univ. of Florida, Gainesville, FL 32611-6250), Christopher Niezrecki (Univ. of Massachusetts Lowell, Lowell, MA 01854-2881, Chris_Niezrecki@uml.edu), and Diedrich O. Beusse (Univ. of Florida, Gainesville, FL 32611-0126)

The West Indian manatee has become endangered partly because of watercraft collisions in Florida’s coastal waterways. Several boater warning systems, based upon manatee vocalizations, have been proposed to reduce the number of collisions. One aspect of the feasibility of an acoustically based system will rely upon the distance in which a manatee vocalization can be detected. The magnitude of environmental noise and manatee vocalizations, as well as the acoustic spreading properties of the habitat will help to estimate the detection range of a manatee. This study surveyed several shallow-water coastal areas in Florida (Crystal River, Cedar Key, and Indian River) which are likely to be inhabited by manatees. Using a chirp signal (1–10 kHz) broadcast by an underwater transducer, it was observed that the acoustic pressure data collected was best represented by the mixed and Lurton spreading models. The overall data obtained from passing a boat directly over the hydrophone was most closely represented by a spherical spreading model. However, for a boat that is approaching, a mixed spreading model is the most appropriate. The mean SPL for boat traffic was measured to be 140 dB and the background environment noise level ranged between 69 and 105 dB.

3:00

4pUWb5. One year of background underwater sound levels in Haro Strait, Puget Sound. Val Veirs (Phys. Dept., Colorado College, Colorado Springs, CO 80903, vveirs@coloradocollege.edu) and Scott Veirs (BeamReach Sustainability Sci. School, Seattle, WA 98115)

Haro Strait, on the west side of San Juan Island, WA, is the home range of the Southern Resident orca whales, a major shipping lane to and from Canada, and a center of private and commercial boating, especially in the summer. Four ITC hydrophones in a near-shore fixed array are used here to localize the underwater vocalizations of Southern Resident orca whales. The system operates 24 hours a day and has a frequency response of 100 Hz to 10 kHz. Background sound levels are automatically characterized by half-hour reports that include: statistics and graphics based on mean sound levels (2-min running arithmetic mean pressure); a histogram of mean sound levels binned by frequency; and 2-s sound samples from maximum background events. Sound levels range from ~90 dB re 1 microPa (quiet conditions) to ~130 dB re 1 microPa when loud commercial ships are passing in the nearby shipping lane or speedboats are passing

close to the hydrophone array. Complete results for one year of continuous monitoring will be presented, segmented by time (season, day of the week, hour in the day), frequency spectrum and dominant noise source class. [Work supported by 35 undergraduate researchers and the Colorado College.]

3:15

4pUWb6. An acoustic modeling study of airgun noise from seismic surveys performed offshore British Columbia. Alexander O. MacGillivray (JASCO Res. Ltd., 2101-4464 Markham St., Victoria, BC, Canada V8Z 7X8) and N. Ross Chapman (Univ. of Victoria, Victoria, BC, Canada V8W 3P6)

A recent numerical modeling study has examined the propagation of underwater noise from potential seismic survey activity in Hecate Strait and Queen Charlotte Sound, Canada. Noise level predictions from this study are based on an integrated modeling approach incorporating an airgun array source model, a broadband transmission loss model, and an environmental model based on high resolution bathymetry, historical CTD casts and geophysical data. Details of the source model, transmission loss model and environmental model are discussed. Selected noise level predictions from the modeling study are presented. Results from this study will aid in evaluating potential environmental impacts of seismic exploration activity on marine ecosystems in British Columbia's offshore region. [Work supported by BC MEM and NSERC.]

3:30

4pUWb7. An integrated acoustic modeling infrastructure for underwater noise impact assessment. David E. Hannay and Roberto G. Racca (JASCO Res. Ltd., 2101-4464 Markham St., Victoria, BC, Canada V8Z 7X8, dave@jasco.com)

When industrial activities such as underwater dredging or seismo-acoustic surveying are planned near ecologically sensitive areas, effective forecasting of the associated sound levels is often necessary to allow effective noise management planning and to satisfy permitting requirements. To this end a comprehensive software application, the Marine Operations Noise Model, was developed to enable the estimation of aggregate noise levels from complex operations involving numerous activities over extended regions. The software architecture includes a specially adapted Parabolic Equation propagation model, a database of field measured spectral source levels from a wide range of industrial vessels, a run module that coordinates modelling and summing of sound from multiple sources, and a GIS interface for the definition of operational layouts and the display of noise level contours on area maps. This application has been used in the planning and regulatory approval process for future pipeline and offshore platform installation activities to take place in the proximity of marine mammal habitats. Its accuracy in this context was validated through an extensive program of acoustic monitoring of similar activities taking place in a nearby location over the course of a prior construction season.

Session 5aAA

Architectural Acoustics: General Topics in Architectural Acoustics

Vincent Valeau, Chair

Univ. of La Rochelle, LEPTAB, Av. M. Crepeau, La Rochelle 17042, Cedex 01, France

Contributed Papers

8:30

5aAA1. An investigation of Columbia College Chicago reverberation chamber. Ryan P. Shannon, Kevin A. Fowler, Jeremy J. Hill, Sean D. Terry (Dept. of Audio Arts & Acoust., Columbia College Chicago, 33 E. Congress Pkwy Ste 601, Chicago, IL 60605, ryanshannon8@hotmail.com), Melinda J. Carney, and Dominique J. Cheenne (Columbia College Chicago, Chicago, IL 60605)

The Audio Arts and Acoustics department at Columbia College in Chicago acquired a new building in 2003. The facility, a former bank, contained an old steel and concrete vault that was converted to a reverberation chamber. The acoustic properties of the space, including reverberation time, modal density, and early reflection maps were analyzed and compared to a computer model. Reflectograms were predicted at various locations and compared to test data acquired with Time Delay Spectrometry (TDS). Polar Energy Time (PET) techniques were also used to identify the location of a single 4×4 sample of foam absorber and the results of the test were compared to the predicted value from the computer model. The results of the tests show that, under its current configuration, the room is usable as a reverberation chamber down to 300 Hz, and that the computer model was able to accurately predict the results from the PET methodology.

8:45

5aAA2. Chronological analysis of architectural and acoustical indices in halls for music performance. Youngmin Kwon and Gary W. Siebein (Architecture Technol. Res. Ctr., School of Architecture, Univ. of Florida, 231 Arch, P.O. Box 115702, Gainesville, FL 32611)

The objective of this study was to identify the changes in architectural and acoustical indices in halls for music performance built in the 18th through the 20th centuries. Seventy one halls were classified in five specific periods from the classical period (1750–1820) to the last two-decade contemporary period (1981–2000) based on chronology in music and architectural acoustics. Architectural indices such as room shape, seating capacity, room volume, balcony configuration, and the like as well as acoustical indices such as RT, EDT, G, C80, IACC, and the like for the halls found in the literature were chronologically tabulated and statistically analyzed to identify trends in architectural and acoustical design for each of the historical periods identified. For example, seating capacity increased from the eighteenth through the early twentieth century. It has decreased in the twentieth century while the room volume per seat has increased. This results in longer occupied reverberation times, higher G values and lower C80 values for contemporary halls compared to those built earlier in the century. Acoustical indices were in general found to be lower during the early twentieth century relative to other periods and to have increased since then.

9:00

5aAA3. Acoustic design for Teatro Sesc Pinheiros. Jose Augusto Nepomuceno and Anthony McCreery (Acustica & Sonica, Fradique Coutinho, 955 sala 01, Sao Paulo, Brasil)

Sesc is a non-governmental social organization to promote culture, sport and education in Brasil. Among its facilities there is the new Sesc-Pinheiros center with a multipurpose theater seating 1000. Due to changes

in the planned use of the theater during construction, Acustica & Sonica was hired by the contractor Mendes Junior to adjust and improve the room acoustics and noise control design of the theater. The original design was modified by adding adjustable acoustical systems, changing in the ceiling and balcony shaping, and revising the specification of all finish materials. The room acoustics modifications followed the recommendations of Akustiks, acoustical consultants from South Norwalk, Connecticut, USA. A concert shell was designed and fabricated to allow the use of orchestral music. An open shell concept was used. The new hall has received excellent critics from pianists, conductors and players, and is rated among the best in the country. Acoustical measurements were performed. The measured values have been compared with the results of a computer model carried out by Acustica & Sonica for this project and they are consistent. [The authors would like to thank Mendes Junior, Sesc, Bengt-Inge Dalenbeck and Christopher Blair of Akustiks for their support.]

9:15

5aAA4. Characterization and prediction of low frequency sound in workrooms. Galen Wong and Murray Hodgson (UBC Acoust. and Noise Res. Group, 3rd Fl., 2206 East Mall, Vancouver, BC, Canada V6T 1Z3)

Low-frequency noise is of concern in industrial settings since workers are subjected to long-term exposures of low-frequency noise caused by machinery which can adversely affect human health and well-being. Passive noise-control methods are usually not cost-effective. Active noise control, on the other hand, is most effective at low frequencies. As part of a project investigating active noise control of low-frequency noise in these settings, and how to predict its benefit, measurements were performed on the propagation of low-frequency noise in a workroom, with and without fittings the obstacles in the room to understand their effects and importance in prediction models. Measurements performed in a real-world workroom, as well as in a scale model workroom show that fittings alter the sound field in an unpredictable manner. With low-frequency octave band filtered noise output, changes in the sound pressure level of 10 dB or more were measured in both cases. Prediction of low-frequency noise in workrooms was performed using an image-phase and a modal model—which do not account for the effects of fittings. Predictions were compared to measurements in full-scale and scale-model workrooms with pure tone outputs. So far it has proven difficult to accurately model low-frequency effects in large workrooms.

9:30

5aAA5. Evaluation and optimization of acoustical environments in eating establishments. Zohreh Razavi and Murray Hodgson (UBC Acoust. & Noise Res. Group, SOEH, 3rd Fl., 2206 East Mall, Vancouver, BC, Canada V6T 1Z3)

Complaints from hard-of-hearing faculty members at UBC about communicating verbally in one of the upscale restaurants on the campus led to a study of acoustical environments in eating establishments and how to optimize them. Speech intelligibility and speech privacy are important considerations in the design of eating establishments; however, they are commonly overlooked. In this preliminary research, three bistros were considered. Physical measurements were done and associated acceptability criteria applied to evaluate the environments. The noise exposures of em-

ployees and customers were measured and compared to maximum permitted occupational limits. Worker noise exposures were in the range 80–84 dBA. Customers were often exposed to levels over 75 dBA; especially at lunchtime it exceeded 80 dBA. The CATT room-acoustical prediction model was used to study the physical and acoustical factors that affect the acoustical environments in eating establishments and how optimize them. The effect of facility modifications, including the use of sound-absorbing wall panels, lowered and improved acoustical ceilings, and partial barriers between tables were predicted.

9:45

5aAA6. A new method for measuring the speech security of meeting rooms. John S. Bradley and Bradford N. Gover (Inst. for Res. in Construction, Natl. Res. Council, Montreal Rd., Ottawa, Canada K1A 0R6)

This paper describes a proposed new method for measuring the speech security of offices and meeting rooms. It is intended to assess the risk of someone who is talking in the meeting room being overheard at points outside the room, close to the outside boundaries of the room. Previous work has derived frequency-weighted speech-to-noise type measures that are well related to the perceived intelligibility and audibility of the transmitted speech. The proposed new method measures the attenuation between average levels in the meeting room and received levels at points 0.25 m from the outside boundaries of the meeting room in adjacent spaces. The average source room levels are representative of all possible talker location. The receiving points, 0.25 m from the room boundaries, can assess expected privacy at critical listening locations with minimal effect of the acoustical properties of the adjacent spaces. New measurements in 11 meeting rooms have been used to evaluate the influence of several parameters on the accuracy of the measured degree of speech security. These include: the type and number of sound sources used, and the number and location of microphones in the meeting room and the adjacent spaces. Details of the analyses will be presented.

10:00–10:15 Break

10:15

5aAA7. A simplified method to estimate the free path length variance. Yan Zhang, Godfried Augenbroe, Ruchi Choudhary (College of Architecture), and Brani Vidakovic (Georgia Inst. of Technol., Atlanta, GA 30332)

Kutruff derived an analytical solution for the prediction of reverberation time in spaces with arbitrary shape. It is the most popular analytical equation that contains shape factors. However, in this equation the free path length variance remains as unknown parameter. Kutruff suggested the implementation of Monte-Carlo simulation to determine this value for each unique shape. In practice this method is not significantly simpler than a full ray tracing simulation. Fortunately, using methods from the probability field, the free path length variance does have an analytical solution, but this takes such a complicated form that it is not convenient to use. This article treats the development of a simplified method to estimate the free path length variance without losing accuracy when applied to concert halls. A simplified regression model is developed for rectangular shapes based on the analytical solution. This simple model explains 99.8% variance for any rectangular shape up to 1:10:10. Secondly, for arbitrary shapes, a simplified model is proposed by including several additional simple variables. This model is validated against simulation results. It will be shown that this simplified method not only can be used in Kutruff's equation, but also reveals the significant shape-related parameters that influence the reverberation time.

10:30

5aAA8. Acoustic modeling in rectangular rooms with impedance using the finite element method with the Dirichlet-to-Neumann map. Yusuke Naka (Dept. of Aerosp. and Mech. Eng., Boston Univ., 677 Beacon St., Boston, MA 02215, ynaka@bu.edu), Assad A. Oberai and Barbara G. Shinn-Cunningham (Boston Univ., Boston, MA 02215)

Computational models of head-related transfer functions (HRTFs) are useful in investigating the effects of echoes and reverberations in enclosures. These models may be computed at relatively low cost by geometric methods, such as the image source method. However, geometric methods typically ignore several important physical effects, such as diffraction, which effect the fidelity of the resulting HRTF. On the other hand, methods based on solving the wave equation, such as the finite element method, include these effects but tend to be computationally expensive. This study represents a Dirichlet-to-Neumann (DtN) map which significantly reduces the costs associated with using the the finite element method for computing HRTFs in a rectangular room, by analytically eliminating empty regions of the room. The DtN map for rooms with realistic impedance boundary conditions is developed. This work represents an extension of our previous approach for sound-hard rooms [Y. Naka, A.A. Oberai, and B.G. Shinn-Cunningham, Proc. 18th International Congress on Acoustics, Vol. IV, pp. 2477–2480 (2004)]. [Work supported by AFOSR.]

10:45

5aAA9. Analysis of diffuse broadband sound fields in enclosures by decomposition in powers of an absorption parameter. Donald B. Bliss, Jerry W. Rouse, and Linda P. Franzoni (Mech. Eng. and Mater. Sci., Duke Univ., Durham, NC 27708, dbb@duke.edu)

A novel analysis for steady-state or time-dependent broadband diffuse sound fields in enclosures is developed. Beginning with a formulation utilizing broadband intensity boundary sources, the strength of these wall sources is expanded in a series in powers of an absorption parameter, thereby giving a separate boundary integral problem for each power. The first problem has a uniform interior field level proportional to the reciprocal of the absorption parameter, as expected. The second problem gives a mean-square pressure distribution that is independent of the absorption parameter and is primarily responsible for the spatial variation of the reverberant field. This problem depends on the location of sources and the spatial distribution of absorption, but not absorption level. Additional problems proceed at integer powers of the absorption parameter, but are essentially higher order corrections to the spatial variation. The important second problem and the higher order problems are easily solved numerically, but closed form approximate solutions based on one iteration step are simple and accurate. Solutions obtained by retaining only the first couple of terms in the series expansion are compared to complete solutions obtained by another approach, and good agreement is shown between the two methods.

11:00

5aAA10. Finite-sized sources in the presence of boundaries. John R. MacGillivray (Red Tail Hawk Corp., 111 E. Seneca St., Ithaca, NY 14850) and Victor W. Sparrow (Penn State, University Park, PA 16802)

Many methods of auralization convolve a source signal (e.g., cello recorded in an anechoic room) with a room's impulse response (which has been computed using method of images, ray tracing, etc.). Many instruments are finite-sized sources because they produce music having frequencies where the product of the wavenumber and the instrument's characteristic length is not small. Sound produced by a finite-sized source in the presence of boundaries can include scattering and diffraction, resulting from the presence of the source in its own field. These effects are not accounted for by the auralization types mentioned above. A geometrically simple example of a finite-sized pulsating sphere in the presence of a rigid infinite boundary is solved using the translational addition theorem for spherical wave functions (TATSWF). Using TATSWF, the original problem is solved by replacing the rigid infinite wall with an image of the finite-sized sphere. This is a surprisingly complicated problem to solve,

given the simple geometry, and serves to illustrate how a source can perturb its field when near a boundary. Examples are presented for which significant changes in the pressure magnitude occur. [Work supported by the Applied Research Laboratory, Penn State.]

11:15

5aAA11. Subjective impression of differences in realism, source width, and orientation between auralizations created from multi-channel anechoic recordings. Michelle C. Vigeant, Lily M. Wang (Architectural Eng. Prog., Univ. of Nebraska–Lincoln, Peter Kiewit Inst., 1110 S. 67th St., Omaha, NE 68182-0681, mvigeant@unlnotes.unl.edu), and Jens Holger Rindel (Tech. Univ. of Denmark, DK-2800 Kgs. Lyngby, Denmark)

Auralizations can be very useful in the design of performing arts spaces. One of the fundamental modeling inputs to create auralizations is the source directivity. Standard methods involve inputting the measured source directivity, calculating the impulse response (IR) and convolving it with a single channel anechoic recording. This paper focuses on an alternative method of modeling source directivity which involves multi-channel anechoic recordings to create auralizations. Subjective tests were conducted comparing auralizations made with one, four and thirteen channels for differences in realism and source width. Auralizations were made using three different types of musical instruments: woodwinds (flute), brass (trombone) and strings (violin). Subjects were asked to rate each musical track on a seven-point scale for the degree of realism and source width. An analysis of variance (ANOVA) was carried out to determine the differences between the number of channels and the effect of instrument. A second test was conducted to assess the degree of difficulty in detecting source orientation (facing the audience or facing the stage wall) depending on the number of channels (one, four or thirteen) and the amount of absorption in the room. [Work supported by the National Science Foundation.]

11:30

5aAA12. Sampling methods for decay time evaluation in acoustically coupled spaces. Tomislav Jasa, Ning Xiang, and Mendel Kleiner (School of Architecture and Dept. of Electrical, Computer, and Systems Eng., Rensselaer Polytechnic Inst., Troy, NY 12180)

This paper applies the methods of Bayesian inference to the estimation of decay times in coupled rooms. Previous papers [N. Xiang and P. M.

Goggans, J. *Acoust. Soc. Am.* **110**, 1415–1424 (2001); **113**, 2685–2697 (2003)] developed a solution method to estimate the decay times and the number of decay modes in terms of measured Schroeder's decay functions. This paper extends the previous work by using statistical sampling methods to efficiently determine the decay times along with error estimates and evaluate the "Bayesian evidence" term used in determining the number of decay modes. This paper discusses the implemented methods together with the previous work to solve the problem of decay time estimation as well as determining the number of decay modes in acoustically coupled rooms.

11:45

5aAA13. Experimental validation of a diffusion equation-based modeling of the sound field in coupled rooms. Alexis Billon, Vincent Valeau (LEPTAB Univ. of La Rochelle, Av. M. Crepeau 17042, La Rochelle Cedex 01, France, abillon@univ-lr.fr), Judicael Picaut (LCPC Nantes ESAR Rte. de Bouaye–BP 4129 44341, Bouguenais Cedex, France), and Anas Sakout (LEPTAB, La Rochelle, France)

Sound modeling in coupled rooms (i.e., two acoustically coupled rooms separated by an open area) has attracted considerable attention in the past. However accurate and operational models are still needed, principally when three or more rooms are coupled. In recent papers, a diffusion equation-based model has been applied to unusual room shapes. For the coupled rooms geometry, this diffusion model has been validated successfully by comparison with the classical statistical theory in a parametrical study of the coupling parameters [Billon *et al.*, *J. Acoust. Soc. Am.* **116**, 2553 (2004)]. In the present work, the diffusion model results are validated by means of a comparison with experimental results, both in terms of sound attenuation and reverberation time. A comparison is also provided with results given by the statistical theory and a ray tracing program. For this purpose, experiments have been conducted in two coupled classrooms with two different sound source locations. The results show a very good agreement between the diffusion model and the experiments. Conversely, the statistical model is not valid for modeling accurately the sound field distribution and decay in both coupled rooms. At last, the diffusion model runs much faster than the ray tracing program.

Session 5aAB**Animal Bioacoustics and ASA Committee on Standards: Behavioral Audiometric Methods in Animal Bioacoustics: The Search for Standards I**

Edward J. Walsh, Cochair

Boys Town National Research Hospital, Omaha, NB 68131

Ann. E. Bowles, Cochair

*Hubbs-Sea World Research Inst., 2595 Ingraham St., San Diego, CA 92109***Chair's Introduction—8:50*****Invited Papers*****9:00****5aAB1. Common hearing functions among mammals and birds.** Richard R. Fay (Parmly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, rfay@luc.edu)

Among many mammals and birds, frequency-place cochlear maps have a similar form and differ primarily with respect to the length of the cochlea or basilar papilla, and with respect to the frequency range of hearing. This is remarkable, especially given the differences in anatomical organization and evolutionary history between mammals and birds. This is a striking example of convergent or parallel evolution. With few exceptions, the cochleae of mammals and birds are indeed scale models of one another. One of the primary functions of the frequency-place map is in frequency analysis. Whether this function is defined by frequency discrimination thresholds, critical masking ratios, critical bands or auditory filters in general, frequency analytic performance, behaviorally-defined, can be laid out on the basilar membrane and a critical basilar membrane distance can be associated with a given performance criterion. In general, these critical distances differ according to the function (e.g., frequency delta F s versus critical bandwidths), but are remarkably similar functions of frequency within and between species (they tend to be constant). It seems likely that evolution has selected the consequences of frequency analytic mechanisms for perception rather than the mechanisms themselves. [Work supported by NIH.]

9:20**5aAB2. How do you know what an animal can hear?** Henry E. Heffner (Dept. of Psych., Univ. of Toledo, Toledo, OH 43606)

The comparative study of hearing is based on the ability to obtain valid behavioral measures of hearing in different species. This requires careful generation and measurement of sound, a behavioral task and reinforcer appropriate for the species, and a comparable definition of threshold. For audiograms, it is important to generate artifact-free pure tones and test with the animal's head fixed in a free-field sound field. For sound localization, a two-point discrimination with the animals head fixed in the sound field is commonly used. For all discriminations, it is important to obtain data for stimuli that are below threshold to ensure that an animal is not using extraneous cues. Descriptions of techniques should be sufficiently detailed to allow other researchers to replicate them; new techniques can be calibrated by replicating the results of others. Behavioral procedures that work well with one species may not elicit optimal performance from others. Although performance should be corrected for false positives, signal detection measures are difficult to interpret and cannot compensate for the failure to keep an animal under stimulus control. In general, valid results are obtained by following good scientific practice.

9:40**5aAB3. The critical band and the critical ratio.** William A. Yost and William Shofner (Parmly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626)

The critical band is arguably one of the most important psychophysical descriptors of auditory function. The critical ratio is often used in animal psychophysical research to estimate critical bandwidth because it is easier to obtain than estimates of critical bandwidth. However, in most cases the underlying assumptions required to use the critical ratio as an estimate of critical bandwidth are rarely met. The critical bandwidths for chinchilla and guinea pigs are similar to those obtained for human subjects, while the critical ratios are considerably higher, an observation that violates the use of the critical ratio as an estimate of critical bandwidth. This paper will describe the assumptions necessary for the critical ratio to estimate critical bandwidth. We will describe data on critical bands and critical ratios for chinchilla and guinea pigs, and we will provide evidence that the chinchilla appears to be a wideband processor of the signal in critical-ratio experiments. The inability of the chinchilla to process the signal in its narrow critical band in the critical-ratio psychophysical procedure leads to the large critical ratios. We recommend that the critical ratio not be used to estimate critical bandwidth. [Work supported by NIDCD grants to Dr. Yost and to Dr. Shofner.]

10:00–10:10 Break

10:10

5aAB4. Estimating loudness in animals. Micheal L. Dent (Dept. of Psych., 361 Park Hall, Univ. at Buffalo-SUNY, Buffalo, NY 14260, mdent@buffalo.edu)

Although auditory acuity experiments in animals are valuable, this is not how animals typically encounter real world sounds. It is also important to know how animals respond to stimuli at more “realistic” intensities. The perception of above-threshold level sounds has been well documented in humans, but is much less studied in animals, where acoustic communication is important for survival. In humans, equal loudness contours are used to show that loudness does not increase with intensity equally at all frequencies. Loudness is a subjective experience, however, and much harder to measure in animals. Reaction time is one untrained response that can be used to gauge how different or similar two stimuli are: high intensities yield short and low intensities yield long reaction times. Comparing reaction times across frequencies (equal latency contours) in birds and mammals show results that are similar to equal loudness contours obtained in humans. At low SPLs equal-latency contours closely parallel threshold curves, while at high SPLs the contours flatten and all frequencies are perceived as being about the same loudness. These results in animals should be considered when thinking about aspects of acoustic communication such as sound transmission, vocal signals designs, and sound-attenuating properties of the environment.

10:30

5aAB5. Animal behavioral psychoacoustics: Issues related to methodology and interpretation. David Kastak, Ronald Schusterman, Brandon Southall, Marla Holt (UCSC Long Marine Lab., Santa Cruz, CA 95060), and Colleen Kastak (UCSC Long Marine Lab., Santa Cruz, CA 95060)

A brief survey of the literature in animal behavioral psychophysics shows that researchers use numerous methods to obtain information on sound detection, discrimination, and identification. Behavioral methodology in animal psychoacoustics includes both classical and operant conditioning, go/no-go or multiple alternative forced-choice tasks, and various methods of estimating detection and discrimination thresholds. Recent emphasis on comparing data across subjects, species, and media (e.g., hearing in air versus water), as well as investigations into the effects of age, noise, and ototoxins on auditory perception highlight the need for methodological standardization. In this paper we will discuss several important issues related to behavioral audiometrics, focusing primarily on marine mammals. These issues include variability among species, individual subjects, and laboratories; experimental concerns such as time constraints; adaptive versus non-adaptive psychophysical methodology and threshold reliability; signal detection theory versus threshold models of audition and the search for unbiased estimates of auditory performance; and measurement and interpretation of subject response bias. Standards for animal psychoacoustic methodology should be sensitive to each of these factors.

10:50

5aAB6. Body size and assessment of auditory function: A comparative conundrum. Edward J. Walsh, JoAnn McGee (Boys Town Natl. Res. Hospital, Omaha, NE 68131), John Rosowski, and William Peake (Harvard Med. School, Boston, MA 02115)

One goal of the bioacoustics community is to compare auditory function among species representing the entire animal kingdom, including terrestrial mammals. As an alternative to behavioral measures, it is frequently necessary, and/or desirable, to assess auditory function using electrophysiological approaches. Body size is an important factor that can effect the distribution and amplitude of evoked brain potentials (EP) measured from the surface of the head and the ranges of body mass and size within *Mammalia* are extensive. Consequently, the development of comparison protocols must include consideration of factors affected by size differences, e.g., the distance between EP generators and recording electrodes and the thickness of the skull. Ultimately, these factors, along with the acoustical character of the recording environment itself, affect acquired signal-to-noise ratios (SNR). In this context it is notable that the SNR associated with large animals are reduced relative to those observed in smaller animals, making the comparison of results from one species to another complex. This procedural challenge is further complicated by the requirement to acquire data efficiently and rapidly in recording environments that are non-optimal from an acoustic perspective. These issues will be addressed by considering auditory brainstem responses in tigers, bobcats, manul, sandcats and rodents.

11:10

5aAB7. Psychometric data and standards for the estimation of noise exposure for animals. Ann E. Bowles (Hubbs-SeaWorld Res. Inst., 2595 Ingraham St., San Diego, CA 92109, annb1@san.rr.com)

ASA standards are used in the estimation of noise exposure for humans. The approaches pioneered for humans can be used as a model for animal standards, but a number of differences must be considered. First, animal standards will normally be applied across multiple taxa rather than to a single, monotypic species. Thus, it will be essential to find defensible methods for generalizing across taxa. Second, samples of subjects and measurement conditions are often inadequate in animal studies (e.g., measurements on a single animal, noisy testing conditions), but may be the only data available. Therefore, standards are needed for specifying the limitations of various data sources. Third, taxa may have very different psychoacoustic capabilities (e.g., best sensitivity, temporal integration time, and loudness perception). Therefore, comparative measures will be essential. For example, while weighting functions for humans are standardized to zero at 1 kHz, there are advantages to developing weighting functions for animals with an appropriate correction for taxon-specific sensitivity. Although the differences between standards for animals and humans represent a significant challenge, the development process and research in support of animal standards will yield valuable new perspectives that could be applied to humans in the future.

Session 5aBBa

Biomedical Ultrasound/Bioresponse to Vibration: Hemostasis and Bleeding Detection

Tyrone M. Porter, Chair

Univ. of Cincinnati, Biomedical Engineering, 231 Albert Sabin Way, Cincinnati, OH 45267-0761

Contributed Papers

8:00

5aBBa1. Vector-Doppler ultrasound for the detection of internal bleeding. Bryan W. Cunitz, Peter J. Kaczkowski, and Andrew A. Brayman (Ctr. for Industrial and Med. Ultrasound, Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

A vector Doppler (VDop) ultrasound system uses a transmitter and a spatially separated pair of receivers to measure bistatic scattering from blood. VDop has two principal advantages over color-flow Doppler in identifying internal bleeding: (1) measures flow direction, and thus absolute magnitude of flow velocity (2) does not require special orientation to detect and measure flow, thus can measure flows perpendicular to the transmitter. Our hypothesis is that real-time flow direction and magnitude can be used to detect and characterize internal bleeding. A real-time vector Doppler system has been built and tested *in vitro*. The system is capable of measuring flow magnitude and direction up to 145 cm/s at a depth of 3.6 cm at a processing rate of 10 Hz. Accuracy was measured using a calibrated moving string phantom and the system performs well within a useful range. A blood flow phantom was developed to mimic arterial flow into an open cavity as well as into tissue and replicate both pulsatile flow as well as the energy storage due to vascular elasticity. Flow signature data is gathered under conditions of normal branching flow, and vessel breach. The talk will describe the VDop system and the flow phantom and summarize results.

8:15

5aBBa2. Color-Doppler guided high intensity focused ultrasound for hemorrhage control. Vesna Zderic, Brian Rabkin, Lawrence Crum, and Shahram Vaezy (Appl. Phys. Lab. and Bioengineering, Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, vesna@u.washington.edu)

To determine efficacy of high intensity focused ultrasound (HIFU) in occlusion of pelvic vessels a 3.2 MHz HIFU transducer was synchronized with color-Doppler ultrasound imaging for real-time visualization of flow within blood vessels during HIFU therapy. HIFU was applied to pig and rabbit pelvic vessels *in vivo*, both transcutaneously and with skin removed. The *in situ* focal intensity was 4000 W/cm² on average. Vessel occlusion was confirmed by color or audio Doppler, and gross and histological observations. In rabbits, five out of 10 femoral arteries (diameter of 2 mm) were occluded after 30–60 s of HIFU application. The average blood flow reduction of 40% was observed in the remaining arteries. In pigs, out of 7 treated superficial femoral arteries (2 mm in diameter), 4 were occluded, one had 80% blood flow reduction, and 2 were patent. In addition, 3 out of 4 superficial femoral arteries, punctured with 18 gauge needle, were occluded after 60–90 s of HIFU application. Larger vessels (diameter of 4 mm) were patent after HIFU treatment. Doppler-guided HIFU has potential application in occlusion of injured pelvic vessels similar to angiographic embolization.

8:30

5aBBa3. A pulsatile flow phantom for image-guided high-intensity focused ultrasound studies on blood vessels. Robyn Greaby and Shahram Vaezy (Ctr. for Medical and Industrial Ultrasound, APL, Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698, rgreaby@u.washington.edu)

A pulsatile flow phantom has been developed for controlled studies of acoustic hemostasis and high intensity focused ultrasound (HIFU) effects on blood vessels. The flow phantom consists of an excised carotid artery attached to a pulsatile pump and embedded in an optically and acoustically transparent gel to create an *ex vivo* model of a human artery. The artery was punctured with a needle to simulate percutaneous vascular injury. A HIFU transducer with a focal distance of 6 cm and a frequency of 3.33 MHz was used to treat the puncture. *B*-mode and power-Doppler ultrasound were used to locate the site of the puncture, target the HIFU focus, and monitor treatment. Also, the effects of vascular flow on HIFU lesions were studied by treating vessels in the phantom for different times at a variety of flows. In both studies, histology was done on the artery. In nine trials, HIFU was able to provide complete hemostasis in 55 ± 31 s. It is feasible to use the flow phantom to study acoustic hemostasis of blood vessels. Histology shows that the flow in the phantom appears to diminish the vascular damage from HIFU.

8:45

5aBBa4. High throughput high intensity focused ultrasound (HIFU) treatment for tissue necrosis. Vesna Zderic, Jessica Foley, Sean Burgess, and Shahram Vaezy (Appl. Phys. Lab. and Bioengineering, Univ. of Washington, 1013 NE 40th St., Seattle WA 98105, vesna@u.washington.edu)

To increase HIFU throughput, a HIFU transducer (diameter of 7 cm, focal length of 6 cm, operating frequency of 3.4 MHz), coupled to an imaging probe (ICT 7-4, Sonosite), was driven by 100 W driving unit with 500 W amplifier. A water pillow connected to a circulation pump with degassed water provided transducer coupling and cooling. Input electrical power was 400 W, corresponding to focal intensity of 51 300 W/cm² (in water). HIFU was applied transcutaneously in porcine thigh muscle *in vivo*, and intraoperatively in liver hilum, for 30–60 s. After 30 s of treatment, muscle lesions had diameter of 2.1 ± 0.4 cm and length of 2.0 ± 0.4 cm. After 60 s of treatment, muscle lesions had diameter of 3.1 ± 0.9 cm and length of 3.1 ± 1.0 cm. In few cases of severe skin burns, the lesions were formed immediately under the skin and were shallow (~ 1 cm). Liver lesions had diameter of 2.1 ± 0.3 cm and length of 2.3 ± 0.6 cm. In comparison, after the 30 s treatment with our standard HIFU device (focal intensity of 27 000 W/cm², in water) the muscle lesions had diameter of 1.3 ± 0.6 cm and length of 1.1 ± 0.6 cm. High power HIFU device can quickly coagulate large tissue volumes.

5aBBa5. Image-guided high intensity focused ultrasound annular array for hemostasis and tumor treatment. Vesna Zderic, Robert Held, Thuc Nguyen, and Shahram Vaezy (Appl. Phys. Lab. and Bioengineering, Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, vesna@u.washington.edu)

To develop and characterize an ultrasound-guided high intensity focused ultrasound (HIFU) array, an 11-element annular phased array (aperture of 3.5×6.0 cm, focal depth of 3.0–6.0 cm, frequency of 3 MHz) was coupled to an imaging probe (C9-5, Philips). LabView software was developed to control driving electronics and image guidance. Radiation force balance measurements, Schlieren imaging, and hydrophone field mapping were performed. Lesions were produced in gel phantoms, and *ex vivo* porcine liver and human cancerous uterus. The lesions were formed beginning at a focal depth of 6.0 cm and moving by 0.5 cm increments to 3.0 cm, and vice versa, with the overall treatment time of 2 min. The transducer had efficiency of 38%, with intensities of up to 6200 W/cm^2 at the natural focus of 5 cm, in water. The 6 dB focal area varied from 0.4 mm^2 (at 3 cm) to 1.5 mm^2 (at 6 cm). The 3 to 6 cm tissue lesions were 2.7 ± 0.5 cm in length, compared to 4.1 ± 0.3 cm for the 6 to 3 cm lesions. The average lesion width was 1 cm. Image-guided HIFU array may enable treatment of large tumor volumes and hemorrhage control at different depths inside the body.

9:15

5aBBa6. Image guided acoustic hemostasis. Sean Burgess (Dept. of Bioengineering, Univ. of Washington, Box 355640, 1013 NE 40th St., Seattle, WA 98105, sburgess@u.washington.edu), Vesna Zderic (Univ. of Washington), and Shahram Vaezy (Univ. of Washington)

Previous studies have shown that high intensity focused ultrasound (HIFU) can successfully control visible bleeding from solid organ injuries. This study investigates the ability of ultrasound image-guided HIFU to arrest occult hemorrhaging in the posterior liver parenchyma using a pig model. The image-guided HIFU device consisted of an intraoperative imaging probe and a spherically-curved HIFU transducer with focal length of 3.5 cm, frequency of 3.23 MHz, and focal acoustic intensity of 2350 W/cm^2 . A total of 19 incisions (14 HIFU-treated and 5 control incisions) were made in five pig livers. The incisions were 30 mm long and 7 mm deep with HIFU application occurring within 20 s of making an incision. Hemostasis was achieved in all treated incisions after a mean \pm SD of 65 ± 15 s of HIFU application. The mean blood loss rate of the control incisions initially and after seven minutes was 0.268 and 0.231 mL/s, respectively. Subsequent histological analysis showed coagulative necrosis of liver tissue around the incision which appeared to be responsible for hemostasis. Ultrasound image-guided HIFU offers a promising method for achieving hemostasis in surgical settings in which the hemorrhage site is not accessible.

9:45

5aBBa7. Ultrasound-guided high frequency focused ultrasound neurolysis of peripheral nerves to treat spasticity and pain. Jessica L. Foley, Frank L. Starr III, Carie Frantz, Shahram Vaezy (Ctr. for Industrial and Medical Ultrasound, Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, jlf2@u.washington.edu), Jessica L. Foley, Shahram Vaezy (Univ. of Washington, Seattle, WA 98195), and James W. Little (Univ. of Washington, Seattle, WA 98195)

Spasticity, a complication of central nervous system disorders, signified by uncontrollable muscle contractions, is difficult to treat effectively. The use of ultrasound image-guided high-intensity focused ultrasound (HIFU) to target and suppress the function of the sciatic nerve of rabbits *in vivo*, as a possible treatment of spasticity and pain, will be presented. The image-guided HIFU device included a 3.2-MHz spherically-curved transducer and an intraoperative imaging probe. A focal intensity of $1480\text{--}1850 \text{ W/cm}^2$ was effective in achieving complete conduction block in 100% of 22 nerves with HIFU treatment times of 36 ± 14 s (mean \pm SD). Gross examination showed blanching of the nerve at the treatment site and lesion volumes of $2.8 \pm 1.4 \text{ cm}^3$ encompassing the nerve. Histological examination indicated axonal demyelination and necrosis of Schwann cells as probable mechanisms of nerve block. Long-term studies showed that HIFU intensity of 1930 W/cm^2 , applied to 12 nerves for an average time of 10.5 ± 4.9 s, enabled nerve blocks that remained for at least 7–14 days after HIFU treatment. Histological examination showed degeneration of axons distal to the HIFU treatment site. With accurate localization and targeting of peripheral nerves using ultrasound imaging, HIFU could become a promising tool for the suppression of spasticity and pain.

5aBBa8. Percutaneous high intensity focused ultrasound on pigs. Herman du Plessis (Surgery, 1 Military Hospital, Thaba Tshwane 0143, South Africa) and Shahram Vaezy (Univ. of Washington, Seattle, WA 98195)

Two types of HIFU Handpieces were tested on pigs under general anesthesia. The direct applicator (focal distance 4 cm) was adequate to control bleeding from a liver injury, although direct pressure was effective in a shorter time. The percutaneous application did not work at all, as the fixed focal distance (6 cm) was too deep for the size pigs available, and the diagnostic modality was difficult to integrate with the therapeutic window. We were able to create necrotic lesions in the liver substance, but not in the areas where the injury was situated. More development is needed to build a better, user-friendly application device before HIFU can be used in the clinical situation for control of hemorrhage.

Session 5aBBb

Biomedical Ultrasound/Bioresponse to Vibration and Physical Acoustics: Audible-Frequency Medical Diagnostic Methods, Including Multimode Techniques I

Thomas J. Royston, Cochair

Dept. of Mechanical Engineering, Univ. of Illinois at Chicago, 842 W. Taylor St., Chicago, IL 60607-7022

Hans Pasterkamp, Cochair

Pediatrics Dept., Univ. of Manitoba, 840 Sherbrooke St., Winnipeg, MB R3A 1R9, Canada

Chair's Introduction—8:00

Invited Papers

8:05

5aBBb1. Audible-frequency medical diagnostic methods: Past, present and future. Armen Sarvazyan (Artann Labs., Lambertville, NJ 08530)

From time immemorial, the medical practitioners used audible-frequency diagnostic methods (percussion, pulse tone analysis, lung, heart and abdominal sound listening) for diagnosing and treatment of ailments without much help from other tests. During the last century, despite the fact that the stethoscope became the sign of the physician, auscultatory techniques are becoming outmoded and are often considered of little clinical value in comparison with many other modern diagnostic technologies. But obviously, there is rich information on the normal and diseased states of human organs in audible-frequency signals since characteristic times for many physiologic processes and resonance frequencies of anatomical structures are in that frequency range. In recent years, numerous emerging technologies started to bring new life into the ancient diagnostic approaches using low-frequency signals. Significant part of these new technologies are based on ultrasonic methods such as remote detection of tissue oscillation using the Doppler technique and especially on emerging ultrasonic technologies based on the use of acoustic radiation force. This talk presents an overview of current trends in the development of methods realizing the tremendous untapped medical diagnostic potential of sonic range acoustics.

8:30

5aBBb2. *In vivo* breast vibro-acoustography. Mostafa Fatemi, Azra Alizad, Dana H. Whaley, and James F. Greenleaf (Mayo Clinic College of Medicine, 200 First Ave. SW, Rochester, MN 55905)

Vibro-acoustography is based on audio-frequency harmonic vibrations induced in the object by the radiation force of focused ultrasound. The resulting object vibration produces a hydrophone detectable acoustic emission that is a function of regional dynamic properties of the object at the vibration frequency. The amplitude of the hydrophone signal is mapped into an image that represents the object's vibro-acoustic response to the radiation force by scanning the focused ultrasound across the object. Contrast in Vibro-acoustography images represent several properties of the object, including the dynamic characteristics at the audio frequency and scattering and absorption properties at the ultrasound frequency. A Vibro-acoustography system has been combined with a stereo-tactic x-ray mammography system for *in vivo* breast imaging and has been tested on a number of volunteers. Resulting images show soft tissue structures and microcalcifications within breast with high contrast, high resolution, and no speckles. The results have been verified using x-ray mammograms of the breast. The encouraging results from *in vitro* and *in vivo* experiments suggest further development of vibro-acoustography may lead to a new clinical breast imaging modality. [Work supported in part by a grant from the Susan G. Komen Breast Cancer Foundation. Vibro-acoustography is patented by MF and JFG.]

8:55

5aBBb3. Supersonic shear imaging. Mathias Fink, Mickael Tanter, and Jeremy Bercoff (Laboratoire Ondes et Acoustique, ESPCI, 10 rue Vauquelin, 75005, Paris, France)

There is a great interest for developing *in vivo* quantitative imaging of the shear tissue properties. This can be achieved by using an ultrafast ultrasonic imaging device able to follow in real time the propagation of low frequency transient shear wave in the body. We have developed such an ultrafast scanner that can give up to 5.000 images/s. From the spatio-temporal evolution of the elastic field a local inversion algorithm allows to recover the shear modulus and viscosity map without the limitations induced by diffraction. Transient shear sources are induced by ultrasonic radiation pressure generated by the ultrasonic probe array. The most interesting configuration that induces high amplitude shear waves used a sonic shear source that moves at supersonic velocities. *In vitro* and *in vivo* results will be presented that demonstrate the great interest of supersonic shear imaging.

5aBBb4. Shear wave interferometry, an application of sonoelastography. Clark Z. Wu and Kevin J. Parker (Hopeman 204 ECE Dept., Univ. of Rochester, Rochester, NY 14623, wuzhe@ece.rochester.edu)

Sonoelastography is an ultrasound imaging technique where low-amplitude, low-frequency (LF) vibration is detected and displayed via real-time Doppler techniques. When multiple coherent shear wave sources exist, shear wave interference patterns appear. Two shear wave sources at the same frequency create hyperbolic shaped interference patterns in homogeneous, isotropic elastic media. Shear wave speed can be estimated from the fringe separation and the source frequency. If the two sources are driven at slightly different sinusoidal frequencies, the interference patterns no longer remain stationary. It is proven that the apparent velocity of the fringes is approximately proportional to the local shear wave velocity. With this approach, local shear wave speed in elastic media can be estimated. In addition, with a single shear wave source at frequency f and the ultrasound probe externally vibrated at frequency $f - \Delta f$, a novel type of moving interference between the shear waves and the frame of reference motion is created. The moving interference fringes represent the shape of shear wave wavefronts while traveling at a much slower speed. This approach provides a real-time visualization of shear wave propagation and local wave speed estimation from which local stiffness is inferred. [Work supported by NIH.]

5aBBb5. Magnetic resonance elastography. Richard Ehman and Armando Manduca (Depts. of Radiol. and Bioengineering, Mayo Clinic, Rochester, MN 55905)

The goal of our research is to develop MRI-based methods for assessing the mechanical properties of tissues *in vivo*. We have focused on a novel MRI technique for visualizing propagating acoustic shear waves [Science **269**, 1854–1857 (1995)]. Suitable dynamic shear stress for Magnetic Resonance Elastography (MRE) can be generated by surface drivers, inertial effects, acoustic radiation pressure, or endogenous physiologic mechanisms. The MRE acquisition sequence is capable of visualizing cyclic tissue motion of less than 1 micron in displacement amplitude, with imaging times ranging from 100 ms to several minutes. Inversion algorithms based on continuum mechanics are used to process the acquired data to generate maps of mechanical properties such as depict stiffness, viscosity, attenuation, and anisotropic behavior. We have applied MRE to assess specimens of a variety of tissues, ranging in stiffness from lung to cartilage. Human studies have demonstrated that it is feasible to apply MRE to quantitatively image the mechanical properties of skeletal muscles, gray and white matter in the brain, thyroid, kidney, liver, and skin. Our preliminary clinical studies have to date applied MRE to observe changes in tissue mechanical properties in patients with breast, brain, and thyroid tumors, liver fibrosis, and diffuse diseases of skeletal muscle.

10:10–10:25 Break

Contributed Papers

10:25

5aBBb6. Microscopic dynamic magnetic resonance elastography. Shadi F. Othman, Thomas J. Royston, and Richard L. Magin (Univ. of Illinois at Chicago, 842 W. Taylor St. MC 251, Chicago, IL 60607)

Microscopic magnetic resonance elastography (uMRE) is a high resolution imaging technique for measuring the viscoelastic properties of small synthetic and biological samples. Mechanical shear waves, typically with amplitudes of less than 100 μm and frequencies of 500–600 Hz, are induced using a piezoelectric oscillator directly coupled to the region of interest. By using multiple phase offsets and motion encoding gradients we acquire data that allows the generation of images that depict shear wave motion and the calculation of local values of the tissue viscoelastic properties. Recent MRE investigations are increasingly being conducted at higher spatial resolution to establish histological correlations between elasticity maps and tissue structures; such microscopic MRE studies require stronger static fields, stronger magnetic field gradients, higher performance RF coils, and more compact, higher frequency mechanical actuators. Microscopic MRE experiments were conducted at 11.74 T in a 54 mm diameter vertical bore magnet with a 10 mm diameter \times 25 mm length cylindrical space available for imaging. The field-of-view ranged from 4 to 14 mm. The study was conducted on agarose gel phantoms of different concentrations ranging from 0.25%–1% w. Different biological samples, including frog oocytes and tissue engineered constructs, were also tested.

10:40

5aBBb7. Coupled vibration and sound radiation from a fluid-filled and submerged or embedded vascular tube with internal turbulent flow due to a constriction. Yigit Yazicioglu, Thomas J. Royston, Todd Spohnholtz, Bryn Martin (Univ. of Illinois at Chicago, 842 W. Taylor St. MC 251, Chicago, IL 60607), Francis Loth (Univ. of Illinois at Chicago, Chicago, IL 60607), and Hisham Bassiouny (Univ. of Chicago, Chicago, IL 60637)

The vibration of a thin-walled cylindrical, compliant viscoelastic tube with internal flow and an axisymmetric constriction that results in turbulent fluid flow is studied theoretically and experimentally. Vibration of the tube is considered with internal fluid-coupling only and with coupling to internal flowing fluid and external stagnant fluid or external tissue-like viscoelastic material. The theoretical analysis includes the adaptation of a model for turbulence in the internal fluid and its vibratory excitation of and interaction with the tube wall and surrounding viscoelastic medium. Theoretical predictions are compared with experimental measurements conducted on a flow model system using laser Doppler vibrometry to measure tube vibration and the vibration of the surrounding viscoelastic medium, as well as miniature hydrophones to measure fluid pressure in the tube. Discrepancies between theory and experiment and the coupled nature of the fluid-structure interaction are highlighted. This study is relevant to and may provide further incite into vascular patency and mechanisms of

failure, as well as diagnostics via noninvasive acoustic measurements. [Work supported by NIH EB002511 and HL55296, and Whitaker Foundation BME RG 01-0198.]

10:55

5aBBb8. A multimode sonic and ultrasonic diagnostic imaging system with application to peripheral vascular characterization. Todd Spohnholtz, Thomas J. Royston, Yigit Yazicioglu, Bryn Martin (Univ. of Illinois at Chicago, 842 W. Taylor St. MC 251, Chicago, IL 60607), Francis Loth (Univ. of Illinois at Chicago, Chicago, IL 60607), and Hisham Bassiouny (Univ. of Chicago, Chicago, IL 60637)

Ultrasound (US) medical imaging technology is enhanced by integrating a simultaneous noninvasive audible frequency measurement of biological sounds that could be indicative of pathology. Measurement of naturally-occurring biological acoustic phenomena can augment conventional imaging technology by providing unique information about material structure and system function. Sonic phenomena of diagnostic value are associated with a wide range of biological functions, such as breath sounds, bowel sounds, and vascular bruits. The initial focus of this multimode technology was to provide an improved diagnostic tool for common peripheral vascular complications that result in transitional or turbulent blood flow, such as associated with arteriovenous (AV) grafts and stenoses in common carotid and other arteries due to plaque buildup. We review: (1) the development the multimode system by combining a commercial US system with a novel sonic sensor array and associated instrumentation, and (2) the evaluation of its capability via controlled phantom models of basic subsurface sound sources/structures, as well as simulations of constricted peripheral blood vessels. [Work supported by NIH EB002511 and HL55296, and Whitaker Foundation BME RG 01-0198.]

11:10

5aBBb9. Characterization of a vibro-acoustography system designed to detect kidney stones during lithotripsy. Neil R. Owen, Michael R. Bailey, and Lawrence A. Crum (Appl. Phys. Lab., University of Washington, 1013 NE 40th St., Seattle, WA 98105)

Acoustic properties of a vibro-acoustography system designed to detect kidney stones were measured. Our system was formed with two spherical transducers (10 cm diameter, 20 cm curvature) in degassed water that were confocal and separated by an angle of 30 degrees. They were driven at 1.1 MHz and 1.125 MHz to generate a difference frequency of 25 kHz. The acoustic field was characterized by scattering from a known target, the curved surface of a steel cylinder with 6.4 mm diameter. Waveforms of both the low and high frequency scattered signals were measured for different target locations, different hydrophone locations encircling the target, and different acoustic pressures. Focal dimensions of the -6 db pressure profile measured at 25 kHz and the fundamental were both 3×10 mm, in an elliptical shape, which is highly localized. Scatter amplitude was rather insensitive to hydrophone position when the target was in the focus, quite sensitive to hydrophone position when the target was out of the focus, and increased linearly with the sum of the sources. It is hoped

that this characterization will help improve the understanding of the mechanisms of the targeting technique. [Work supported by NIH grants DK43881 and DK55674, and NSBRI grant SMS00203.]

11:25

5aBBb10. 3D steady-state ultrasound-elastography. Ralph Sinkus, Jeremy Bercoff, Mickael Tanter, and Mathias Fink (Laboratoire Ondes et Acoustique, 10 rue Vauquelin, 75005 Paris, France)

One of the current main limitations of Elastography is the lack of access to the full displacement field within a volume. Standard ultrasound techniques provide good motion estimation only along the ultrasound beam. These data are subsequently used for the reconstruction of elastic parameters assuming that the missing displacement components are negligible. Although feasible under certain circumstances, these assumptions typically do not hold for in-vivo applications. Thus, there is need for a technique to assess the entire displacement field in 3D. Recently, the 1D motion estimation has been extended to the measurement of the 2D displacement field [Tanter *et al.*, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **49**, 1363–1374 (2002)]. This method is utilized for the assessment of the 3D displacement field by rotating and translating the US-transducer around the object of interest. Low-frequency mechanical vibrations (approx. 80 Hz) are coupled into the object from underneath (i.e., from the direction perpendicular to the plane of rotation). The measured displacement field is used to reconstruct within a volume shear elasticity and shear viscosity by local inversion of the wave equation. Contributions from the compressional wave are removed via application of the curl-operator. In-vitro results for phantoms and excised specimen are presented.

11:40

5aBBb11. Feasibility of shear-mode transcranial ultrasound imaging. Greg T. Clement, P. Jason White, and Kullervo Hynynen (Dept. of Radiol., Harvard Med. School, Brigham and Women's Hospital, Boston, MA 02115, gclement@hms.harvard.edu)

Despite ultrasound's potential to provide a low cost method for imaging blood flow and diagnosing certain brain disorders, distortion and low signal to noise ratios caused by the skull have severely limited the use of existing clinical devices, such as transcranial Doppler sonography. Presently we investigate the potential to propagate ultrasound through the skull with reduced distortion and higher signal amplitudes by using high incident angles. In such cases the ultrasound angle of entry is set beyond Snell's critical angle for the longitudinal pressure wave, so that propagation in the bone is purely due to a shear wave. This wave then converts back to a longitudinal acoustic wave in the brain. This conversion from a longitudinal wave (skin) to a shear wave (skull) and again to a longitudinal wave (brain) does not necessarily produce a highly distorted or small-amplitude wave. Basic images and measurements of shear speed-of-sound and attenuation values for *ex vivo* human skull bone will be presented for frequencies between 0.25 MHz and 2 MHz. Similar measurements with porcine samples will also be shown, indicating the large discrepancy between shear characteristics of the two species; the porcine samples showing no detectable shear mode.

Session 5aMU

Musical Acoustics: String Instrument Design and Construction

Thomas D. Rossing, Cochair

Physics Dept., Northern Illinois Univ., De Kalb, IL 60115

Christopher E. Waltham, Cochair

*Dept. of Physics and Astronomy, Univ. of British Columbia, 6224 Agricultural Rd., Vancouver, BC V6T 1Z1, Canada***Invited Papers**

8:15

5aMU1. Acoustical considerations in the design—and re-design—of the violin. Joseph Curtin (Joseph Curtin Studios, 3493 W. Delhi, Ann Arbor, MI 48103, violins@josephcurtinstudios.com)

The violin is a highly evolved instrument which has long resisted significant changes to its design and construction. Still, acoustical research over the past several decades has shed sufficient light on the workings of the violin that makers can now consider non-traditional approaches to their craft in order to optimize the sound, playability, and consistency of their instruments. The work of researchers such as Duenwald, Haines, Hutchins, and Weinreich will be considered in terms of its usefulness as a guide to building better violins.

8:40

5aMU2. The violin octet and bowed string instrument design. George Bissinger (Phys. Dept., East Carolina Univ., Greenville, NC 27858)

Modal analyses were combined with room-averaged acoustic measurements of a complete octet to assess Schellengs fundamental scaling design assumptions: similarity of shape and flat plate scaling. The scaling employed only the two lowest strongly radiating violin resonances, the main air $A0$ and main wood comprised of the first corpus bending modes $B1^-$ and $B1^+$. $A0$ fell below the desired pitch placement ($1.5\times$ lowest string pitch), while the $B1$ complex generally straddled the desired placement at $2.25\times$. Difficulties in properly scaling $A0$ derived partly from an unreliable theory that failed to incorporate $A0$ coupling to $A1$ (first length-wise cavity mode), and partly from inability to incorporate cavity wall compliance. Wall compliance dropped $A1$ into main wood region even though larger instruments were designed successively shorter; the $A1/A0$ frequency ratio rose from ~ 1.5 to ~ 2.0 (smallest to largest). Modern models sensitive to cavity shape predict $A0$ and $A1$ within $\sim 10\%$ over the octet, ranging over $\sim 4.5:1$ in length, $\sim 10:1$ in f -hole area, $\sim 3:1$ in top plate thickness, and $\sim 128:1$ in volume. $A0$ radiates strongly over the octet, while surprisingly $A1$ is the dominant radiator in the main wood region for the large bass even though $A1$ was never included in the scaling.

9:05

5aMU3. The acoustics of hammered dulcimers. David R. Peterson (Dept. of Mathematics, Univ. of Central Arkansas, Conway, AR 72035, DavidP@uca.edu)

The hammered dulcimer, a stringed instrument played with two wooden hammers, probably originated in the Middle East, but it has become part of the musical culture of many countries. In the U. S., the folk revival in the 1970's sparked renewed interest in the hammered dulcimer as a concert instrument. Today, despite some consolidation in the retail market, there are still hundreds of builders, mostly amateurs, who experiment with the basic design. The most important design parameters will be discussed from a practical and acoustical point of view: soundboard size, shape, and composition, internal bracing, bridge shape, string arrangement and composition, hardness of bridge caps, hammer weight and stiffness, instrument resonances due to the unique string splitting and stiffness of the body, and soundboard modes.

9:30

5aMU4. Classical guitar construction: The acoustician's tale. Bernard E. Richardson (School of Phys. and Astron., Cardiff Univ., 5 The Parade, Cardiff CF24 3YB, UK, RichardsonBE@cardiff.ac.uk)

The vast majority of guitars produced today are built according to general principles laid down in the nineteenth century. Nevertheless, the devil is in the detail, and innovative makers constantly reappraise the design and construction of instruments in their endeavors to control quality or to seek "improvement." The maker's approach, necessarily, tends to be pragmatic, but it is one which can be greatly informed by the application of relatively simple acoustical models. This paper will examine various important design aspects—for example choice of materials, body size, strutting, soundboard thickness—introducing the basis for making informed decisions.

9:55

5aMU5. Normal modes of a single mandolin with different bracing patterns installed. David J. Cohen (Cohen Musical Instruments, 9402 Belfort Rd., Richmond, VA 23229) and Thomas D. Rossing (Northern Illinois Univ., DeKalb, IL 60115)

The vibrational modes and sound spectra of a single archtop mandolin constructed with a removable back plate were studied with different bracing patterns installed. The bracing patterns were (i) symmetric and (ii) asymmetric parallel “tone bars,” (iii) X-bracing, and (iv) an asymmetric pattern previously used and studied [D. Cohen and T. D. Rossing, *Can. Aeronautics Space, J.* **4**(2), 48–54 (2000), D. Cohen and T. D. Rossing, *Acoust. Sci. Tech.* **24**, 1–6 (2003)]. The results have been compared with those obtained previously for vintage mandolins [D. Cohen and T. D. Rossing, “Normal modes of different types of pre-1929 mandolins,” presented at the 147th meeting of the Acoustical Society of America, New York, NY (2004)] and for mandolins constructed more recently. As was the case for the different types of vintage mandolins, changing the bracing pattern has little effect on the shapes of the normal modes, but does affect the frequencies of the normal modes. The X- or crossed bracing pattern imparts more cross-grain stiffness than the other bracing patterns, with the result that modes involving cross-grain bending occur at higher frequencies than is the case for the other bracing patterns.

10:20–10:30 Break

10:30

5aMU6. Design aspects of the five string banjo. James Rae (Dept. Physio. and Biomed. Eng., Mayo Clinic College of Medicine, 200 First St. SW, Rochester, MN 55905, rae.james@mayo.edu) and Thomas Rossing (Northern Illinois Univ., DeKalb, IL 60115)

The five string banjo is a stringed instrument that uses a drum head (membrane) to radiate sound. The strings couple to the head through a moveable bridge. Many have a removable resonator whose spacing from the main part of the banjo is adjustable to allow tuning of the internal air cavity. Holographic interferometry demonstrates vibrational modes of the head that have a strong dependence on head tension. Driving point acceleration measurements demonstrate that the ability of the bridge to be vibrated depends on the mass and stiffness of the materials used in its construction. Peak accelerances usually occur between 1800–2000 Hz. Power spectra measurements of the sound show that 99% of the sound is in a frequency range of 147–5200 Hz. Two substantial formants are seen in the power spectra. The first and largest occurs in about the 400–1800 Hz range, the same frequency range where the most substantial head modes are found. The second is from 2000–4000 Hz, a range where the series combination of bridge and head show increasing accelerances by driving point measurements. Measurements of the transient response following a pluck quantify the frequency content of the rising and falling phase.

10:55

5aMU7. Tuning the lower resonances of carved Baltic psalteries by adjusting the areas of the sound holes. Andres Peekna (Innovative Mech., Inc., 5908 N River Bay Rd., Waterford, WI 53185) and Thomas D. Rossing (Northern Illinois Univ., De Kalb, IL 60115)

Previous work has shown that with carved Baltic psalteries it is highly desirable to have a body-resonance with high radiating efficiency between the keynote and the tone above. In many cases, it is also desirable to have a body-resonance close to the low dominant. The frequencies of the two lowest body-resonances can be adjusted, to some extent, by adjusting the areas of the sound holes. We describe two trials. In one case, the instrument started out with no sound holes, and their areas were increased step by step, mostly according to listening trials, but the results were also verified by electronic TV holography. In the second case, the instrument started with excessive sound hole area. In this case, the listening tests first determined the extent of sound holes to be covered by sufficiently stiff material. The findings were subsequently confirmed by electronic TV holography. In both cases, a point was reached below which (smaller areas) the sound quality became less bright and more dull, and above which the sound quality was perceived as good. In each case, an optimum was reached.

11:20

5aMU8. Harp design and construction. Chris Waltham (Dept. of Phys. & Astron., Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z1, waltham@physics.ubc.ca)

The harp is an instrument with a set of plucked strings connected directly to the sound board. Thus the requirement that the sound board be very light for efficient radiation is pitted against the requirement that the strings be tight enough to maintain harmonicity. These factors have determined the evolution of the harp since Sumerian times. As materials and construction methods have improved, strings have become tighter and sound boards lighter. Large harps have total tensions in the range of ten kilonewtons. The material of choice for the sound board, usually spruce, is pushed to the limits of its strength, and even professionally built concert harps tend not to last more than a few decades. The strings themselves are made from a variety of materials, wrapped or plain, to balance the physical requirements of tension and feel, with the acoustical demands of harmonicity and tone. Historical materials such as copper and silk have given way to steel and nylon, although gut still has a niche in the mid-range of a concert harp. These physics and engineering issues will be considered in the context of practical harp construction.

11:45

5aMU9. The motion of harp strings. Gary Chan and Chris Waltham (Dept. of Phys. & Astron., Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z1)

The harp is an instrument with a set of plucked strings that excite the sound board directly, without the medium of a bridge. The strings are

positioned at an acute angle to the plane of the sound board. The quality of the sound produced depends on the motion of the string and its interaction with the resonances of the sound board. The string and sound board motions of small and large harps have been studied using small, fast position sensors. The results are compared to those of a simple non-linear model based on the measured elastic properties of the string materials, and those of the sound board. The implications for the sound production are discussed.

FRIDAY MORNING, 20 MAY 2005

REGENCY F, 8:30 TO 11:55 A.M.

Session 5aNS

Noise and Psychological and Physiological Acoustics: Workshop on Methods for Community Noise and Annoyance Evaluation I

Brigitte Schulte-Fortkamp, Cochair

Technical Univ. Berlin, Inst. of Technical Acoustics, Secr TA 7, Einsteinufer 25, Berlin 10587 Germany

Bennett M. Brooks, Cochair

Brooks Acoustics Corp., 27 Hartford Turnpike, Vernon, CT 06066

Chair's Introduction—8:30

Invited Papers

8:35

5aNS1. Assessing multi-source noise environments with an “equally annoying” exposure summation model. Paul Schomer (Schomer and Assoc., Inc., 2117 Robert Dr., Champaign, IL 61821)

Miedema [J. Acoust. Soc. Am. **116**, 949–957 (2004)] posits an “equally annoying” energy summation model for assessing annoyance in areas exposed to multiple differing noise sources. His “proof” is a theorem. Since his model fits the requirements of the theorem, he states that this “proves” that his model is correct. While it may be true that meeting the conditions of the theorem is a necessary condition, it is not a sufficient condition. From a psycho-acoustical standpoint, the Meidema model is very complex and not really believable. Moreover, it fails to address differences between types of aircraft, types of operation, types of motor vehicle, etc. An “equally annoying” exposure summation model, presented here, is psycho-acoustically much simpler and better satisfies the same theorem conditions because it better satisfies the independence requirement.

8:55

5aNS2. Neural network analysis of soundscape in urban open spaces. Lei Yu and Jian Kang (School of Architecture, Univ. of Sheffield, Western Bank, Sheffield S10 2TN, UK, j.kang@sheffield.ac.uk)

The physical and social environments of open spaces in a city influence people's behavior as well as their perception of the city. Soundscape is an important component of an open urban space, and key aspects for producing a good soundscape include its description, evaluation, and design. In this research, an artificial intelligent system is being developed for the analysis of soundscape evaluation in urban open spaces. The system considers acoustic factors including the level, type and spectrum of sound sources and the reflection effects of an urban space; users social and demographic factors; and other physical and environmental factors including urban morphology, microclimate conditions, and thermal and visual comfort. Since 2001 a large scale social survey has been carried out in England, Switzerland, Italy, Greece, Germany and China. Correspondingly, a database suitable for artificial neural networks (ANN) modeling has been established. Simulating biological brains, ANN are simplified models of the central nervous system, which have the ability to respond to input stimuli and to learn to adapt to the environment. The software Qnet is being used, and initial results suggest that there is a good convergence of using ANN to predict peoples perception of soundscape in urban open spaces.

9:15

5aNS3. Soundscape assessment. Osten Axelsson, Birgitta Berglund, and Mats E. Nilsson (Inst. Environ. Medicine, Karolinska Inst. and Dept. Psych., Stockholm Univ., SE-10691, Stockholm, Sweden)

In order to improve the quality of the soundscape it is necessary to know its descriptive and evaluative properties, and the relationships between these properties. This was explored in a listening experiment with 100 participants (48 women, 52 men; mean age 25.6 years). Each participant scaled 5 out of 50 soundscapes with regard to 116 single verbal attributes, using a visual analogue scale of agreeableness. In addition, acoustical properties of the soundscapes were assessed. A principal component analysis identified

two major evaluative components, labeled Hedonic Tone and Eventfulness. Furthermore, it was found the mere presence of common sound sources, regardless of sound level, correlated significantly with these evaluative components. Technological sounds (e.g., traffic noise) were negatively associated with both Hedonic Tone and Eventfulness, while a positive association was found between Hedonic Tone and sounds of nature (e.g., bird song), and a positive association was found between Eventfulness and human sounds (e.g., human voices). These relationships lead to the hypothesis that introduction of nature and human sounds, in combination with the reduction of technological sounds may improve the quality of soundscapes considerably.

9:35

5aNS4. Assessment of outdoor soundscapes in quiet areas. Mats E. Nilsson and Birgitta Berglund (Inst. of Environ. Medicine, Karolinska Inst. and Dept. of Psych., Stockholm Univ., SE-106 91, Stockholm, Sweden)

Existing quiet outdoor areas should be preserved. Appropriate indicators and limit values are needed, which are grounded in knowledge on positive aspects of soundscapes, such as perceived pleasantness and psychological restoration. For this reason, a questionnaire study was conducted in four green areas close to a major city and in four city parks. Measured equivalent sound levels (LAeq, 15 min) ranged between 42 and 50 dBA in the green areas, and between 49 and 60 dBA in the city parks. Sounds from nature, such as bird song, completely dominated the soundscape in the green areas. The city-park soundscapes were more complex, containing sounds from nature, as well as technological sounds (e.g., traffic noise), and human sounds (e.g., human voices). In general, sounds from nature were perceived as pleasant, technical sounds as annoying, and human sounds as neutral. Between 84 and 100% of the visitors in the green areas assessed the soundscapes as good or very good. The corresponding percentages for the city parks were distinctly lower, between 52 and 65%. The results indicate that the equivalent sound level should be below 50 dBA in order to secure pleasant and restorative outdoor soundscapes in urban areas.

9:55–10:15 Break

10:15

5aNS5. Techniques of analysis and their applicability in the context of community noise. Andre Fiebig and Brigitte Schulte-Fortkamp (ITA-Tech. Univ. Berlin, Einsteinufer 25, D-10587 Berlin, Germany)

There is common consent to the necessity using and combining subjective and objective data for a sufficient understanding of human perception and evaluation of noise. Moreover, a wide discussion about methodology and techniques of analysis with respect to new methods in the context of soundscape research, community noise, and annoyance evaluation is urgently needed. The presentation at hand shows a methodological discussion of qualitative data and its systematic analysis based on explorative approaches applied in two different surveys. For example, by means of the analysis a transfer from verbal evaluations based on habitual language to acoustical descriptors is possible. Furthermore, the principals and steps of the analysis of qualitative data in the context of community noise are presented. Insofar, with regard to an effective and unitized handling, options and possibilities of methods in connection with specific techniques of analysis will be discussed.

10:35

5aNS6. Correlation of airport noise complaint terminology with available noise metrics. Nancy S. Timmerman (Consultant in Acoust. and Noise Control, 25 Upton St., Boston, MA 02118-1609)

In the discussion of airport noise, there is a disconnect between the language used by complainants and the acoustical terminology available to the technical community. It will be shown in what ways the complaint terminology can be described with the available noise metrics and in what ways it cannot. The complaint terminology analyzed comes from the author's personal (previous) experience as noise control officer at Logan International Airport. Examples of some of these are (aircraft noise is): too loud, too late, too early, unending, continuous, at the wrong time, so loud it sets off car alarms, so loud I cannot hear the TV/telephone/radio, so loud it woke me up, and annoying. Terms which are not related to noise are mentioned, but not correlated. The acoustical terminology comes from the international literature and includes the standard noise metrics as well as some more recently proposed. The airport noise metrics discussed include measures of single events (maximum level, SEL, duration, signal-to-noise) and measures of noise exposure (DNL, CNEL, Leq). It is shown that while cumulative noise metrics are useful for planning and comparison, they correlate with complaint terminology less well than single event measures.

10:55

5aNS7. Explorative interviews as a tool for sound evaluation. Stephan Paul (Lab. of Vib. and Acoust., Dept. of Mech. Eng., UFSC, CxP 476 Florianopolis, 88.040-900 SC, Brazil)

Sound and noise evaluation needs to combine physical-measurements of sound with research on the interaction of the sound and the human being. For this kind of research qualitative techniques and especially explorative interviews, with the addressed people are appropriate as they consider the human being acting within this environment. The evaluation process refers to the basic concepts of qualitative research. Different forms of interviews, how to prepare and conduct an interview and not to forget about the discipline under investigation have to be considered. This clearly shows the requirement of transdisciplinarity including knowledge from disciplines like engineering sciences, social sciences and psychology. To obtain the required information on the subject under investigation within an interview the verbal ability of the addressed people has to be considered. The possibilities of transdisciplinary work, check of verbal abilities as well as approaches for basic standardization in interviews and the need for cultural adaptation will be discussed.

11:15

5aNS8. Digital identification of intrusive noise: Pilot study to digitally characterize soundscapes and intrusion. Tim Lavallee (LPES, Inc., 14053 Lawnes Creek Rd., Smithfield, VA 23430), Robert Kull (Parsons, Norfolk, VA 23502), and Brigitte Schulte-Fortkamp (ITA/TU-Berlin, Einsteinufer 25, D-10587 Berlin, Germany)

One of the difficulties with soundscape investigation and the definition of intrusion is the need for an individual to manually identify and log specific acoustical events. Event logging can be labor intensive, costly, and difficult to standardize. If information about physical setting and in situ acoustical signatures can be used to define a given soundscape, event logging could possibly be performed digitally and intrusion could be defined based on the spectral fingerprint of a given soundscape. Two soundscapes with different settings and acoustical signatures were identified. Sound time histories and periodic third octave spectra were measured over a given period. An individual manually logged acoustical events. Independently, algorithms to identify both normal and acoustically intrusive events for the given soundscape were applied to the data. The digitally identified events were compared to the manually taken event log. An evaluation of the results will be presented.

11:35

5aNS9. Evaluation of sound environment characteristics: Comparative study between objective and subjective criteria. Francoise Chartier and Catherine Semidor (GRECO, EAPBx, Domaine de Raba F-33400 Talence, France, catherine.semidor@bordeaux.archi.fr)

The evaluation of urban sound environment quality depends on quantitative information as well as the opinion of city dwellers. In order to underline the relation between objective and subjective points of view, a comparative study was carried out. The subjective survey consisted in listening to several binaural sound recordings of very short extracts from the urban scene, called "sonoscene." During these sessions, the listeners' opinions were noted. The binaural sound recordings are performed in accordance with the "soundwalk" method explained in previous papers [C. Semidor, ASA 2004, ASA 2005]. The same binaural sound recordings were analyzed in form of objective criteria such as Equivalent Sound Level, Loudness, Roughness, etc. This paper deals with the comparison between some of these objective criteria and subjective judgments such as Lack or Presence of traffic noise, Lack or Presence of human activities, Spatial Dimension, Attractiveness or Annoyance, etc. These first results point out some significant correlations between Loudness and Attractiveness for example, according to the nature of the sound sources (traffic, human activity).

FRIDAY MORNING, 20 MAY 2005

REGENCY B, 8:00 TO 10:00 A.M.

Session 5aPAa

Physical Acoustics: Topics in Acoustic Mine Detection

Murray S. Korman, Chair

Physics Dept., U.S. Naval Academy, Annapolis, MD 21402

Contributed Papers

8:00

5aPAa1. Long-range excitation of vibrational response in landmines with Rayleigh and Love waves. Thomas Muir (Natl. Ctr. for Physical Acoust., Univ. Mississippi, One Coliseum Dr., University, MS 38677)

An experiment was conducted in a clay soil at a range of 25 m from two types of electromagnetic shaker sources, operating in a cw pulse mode centered at 100 Hz. The sources were in co-located arrays, one vertically oriented to generate Rayleigh waves and another transversely oriented to generate Love waves. Two cylindrical landmine targets were used: an unloaded (empty) plastic shell, and a metal shell, loaded with paraffin to simulate an explosive. These target types represent extreme variations in what could conceivably be encountered. The targets were instrumented with three-axis seismometers and accelerometers, and were also viewed with a laser Doppler vibrometer (LDV). Data were acquired on both mine cases in various stages of deployment, including sitting upon the ground surface, snugged in to a tight hole in the ground, and completely buried some 5 cm beneath the surface. The responses of the mine cases to Rayleigh and Love wave excitation were measured and compared to data from a reference seismometer deployed nearby, on natural, undisturbed soil. The mine case responses to these seismic excitation fields, under a number of conditions are presented and discussed. [Work supported by the U.S. Marine Corps Systems Command.]

8:15

5aPAa2. Influence of buried objects on the dynamic nonlinearity of inhomogeneous materials. Keith Attenborough, Qin Qin (Dept. of Eng., Univ. of Hull, Hull HU6 7RX, UK), Jonathan Jefferis, and Gary Heald (DSTL, Fort Halstead, Kent, UK)

Donskoy *et al.* [J. Acoust. Soc. Am. **111**, 2705–2714 (2002)] and Korman *et al.* [J. Acoust. Soc. Am. **116**, 3354–3369 (2004)] have shown that the nearby presence of compliant buried objects results in a significant nonlinear response in the spectrum of surface vibration induced by acoustic excitation. The latter have suggested that there is a strong connection with the inherent Nonlinear Mesoscopic Elasticity (NME) of inhomogeneous materials. Laboratory experiments are reported that investigate the comparative NME of granular and fibrous materials and the variation observed between several states of a single material (sand) subject to high sound pressure levels in the frequency range 100 Hz to 2 kHz. Fiberglass and gravel are shown to have strong inherent nonlinearity whereas wet compacted sand has relatively little. In both two-tone and single tone experiments, it is shown that, the presence of buried objects may enhance or reduce the inherent NME of the embedding medium. In cases where the NME is enhanced, compliant buried objects have a greater effect than relatively stiff objects. [Work supported by DSTL, UK.]

5aPAa3. Investigation of the applicability of microphones to nonlinear acoustic mine detection. Douglas Fenneman and Brad Libbey (U.S. Army RDECOM CERDEC Night Vision and Electronic Sensors Directorate, 10221 Burbeck Rd., Fort Belvoir, VA 22060)

The feasibility of using microphones to measure intermodulation effects above a buried mine is under investigation at the U.S. Army RDECOM CERDEC Night Vision and Electronic Sensors Directorate. Acoustic mine detection techniques employ acoustic energy to excite soil and buried mines. The resultant linear and nonlinear response at the surface can then be measured using non-contacting vibrometers. Carrier signal scattering by rough surfaces, however, can limit the measurement accuracy of these vibrometers and, subsequently, successful detection of the landmine. A microphone is proposed as an alternative non-contact sensor specifically for nonlinear acoustic mine detection applications. In this scenario, distinct frequency content facilitates separation of the intermodulation effects at the surface from the acoustic excitation. Experimental results employing intermodulation effects for acoustic mine detection have been reported in the literature [Donskoy *et al.*, *J. Acoust. Soc. Am.* **111**, 2705–2714 (2002)]. Preliminary experimental results demonstrate the ability of microphones to sense pressure radiated from soil in the presence of realistic ground velocities. The applicability of these measurements to practical mine detection systems will also be addressed.

5aPAa4. Application of time-reversal focusing in elastic wave propagation to buried object detection. Pelham D. Norville and Waymond R. Scott, Jr. (School of Elec. and Comput. Eng., Georgia Inst. of Technol., 777 Atlantic Dr., Atlanta, GA 30332-0250)

Time-reversal focusing is a technique for focusing energy both spatially and temporally to a single desired point. First studied in fluid media, additional studies have demonstrated the applicability of time-reversal focusing to solid media and elastic wave propagation. A significant difference between time-reversal and other focusing methods is time-reversal focusing immunity to the variation of wave speeds and the effects of scattering objects within the medium. In the detection of buried objects, this feature is paramount where wave speed variations and clutter are common in the region to be searched. The effectiveness of time-reversal focusing is investigated for a variety of configurations of scattering objects. Fields of uniform objects are examined as well as distributions of randomly shaped clutter, and changes in the propagation medium. Both experimental and three-dimensional numerical results are presented. In previous studies, a phenomenon of super-resolution caused by high-order scattering has been observed where focusing exceeds the diffraction limits of the excitation array. The extent of super-resolution is evaluated for the various configurations of scattering objects in an elastic medium. [Work supported by ARO.]

5aPAa5. Nonlinear acoustic experiments involving landmine detection: Connections with mesoscopic elasticity and slow dynamics in geomaterials, Part II. Murray S. Korman (Phys. Dept., U.S. Naval Acad., Annapolis, MD 21402) and James M. Sabatier (Univ. of Mississippi, University, MS 38677)

In nonlinear acoustic detection, airborne sound at two primary tones, f_1, f_2 (chosen several Hz apart from resonance) insonifies the soil surface over a buried landmine, and due to soil wave interactions with the landmine, a scattered surface profile can be measured by an LDV. Profiles at $f_1, f_2, f_1 - (f_2 - f_1)$ and $f_2 + (f_2 - f_1)$ exhibit a single peak while profiles at $2f_1 - (f_2 - f_1), f_1 + f_2$ and $2f_2 + (f_2 - f_1)$ are attributed to higher order mode shapes. The lowest resonant frequency for a VS 1.6 plastic, inert, anti-tank landmine, buried at 3.6 cm deep is ~ 125 Hz. The “on target” to “off target” contrast ratio, for some of the nonlinearly generated combination tones, is roughly 15–20 dB higher compared to either primary component. Near resonance, the bending (softening) of a family of increasing amplitude tuning curves, involving the surface vibration over the landmine, exhibits a linear relationship between the peak particle velocity

and corresponding frequency. The tuning curves exhibit hysteresis effects. Landmine-soil vibrations exhibit similar characteristics to nonlinear mesoscopic/nanoscale effects that are observed in geomaterials like rocks or granular materials. Nonlinear mechanisms of soil and the soil interacting with the top-plate of the mine case are compared. [Work supported by U.S. Army RDECOM CERDEC, NVESD.]

5aPAa6. Time-dependent finite-element model for optimizing source array element position and excitation for a seismic sonar for buried mine detection. Anthony N. Johnson (CPT, U.S. Army, and Dept. of Mathematics, Naval Postgrad. School, Monterey, CA 93943), Clyde L. Scandrett (Naval Postgrad. School, Monterey, CA 93943), and Steven R. Baker (Naval Postgrad. School, Monterey, CA 93943)

A three-dimensional (3-D) continuum mechanics approach to the development of a time-dependent finite-element model for optimizing the position and excitation of source array elements for use in a seismic sonar to detect buried landmines is presented. Mathematical formulation of the problem consists of the coupling of a system of linear, second order, partial differential equations and related boundary conditions into one single wave equation, from which a composite elastic finite element is derived. The *hp*-adaptive finite-element kernel, ProPHLEX [Altair Engineering, Inc., McKinney, TX], is used to perform the numerical computations. The radiation characteristics of a discrete number of transient seismic sources are analyzed in a linear, isotropic, homogeneous half-space. Results for radial and vertical radiation fields, and for radiated Rayleigh wave strength will be presented for various source configurations. Particular attention will be paid to those configurations which maximize the radiation of unidirectional Rayleigh waves, while suppressing the radiation of unwanted body waves.

5aPAa7. Deduction of ground impedance using level difference measurements. Shahram Taherzadeh (Faculty of Technol., The Open Univ., Milton Keynes MK7 6AA, England) and Keith Attenborough (Univ. of Hull, Hull HU6 7RX, England)

Previously, a numerical method for deducing ground-surface impedance from measurements of complex excess attenuation spectra was reported [*J. Acoust. Soc. Am.* **105**, 2039–2042 (1999)]. Subsequent applications have predicted unrealistic values at low frequencies. Here, we report improved results using two-microphone, transfer-function measurements and discuss its merits as opposed to measuring excess attenuation in regards to a free field. Furthermore, effects of errors in measured parameters on the deduced impedance are discussed also.

5aPAa8. Comparison of two mapping methods for computing sound propagation over irregular terrain. Xiao Di (Appl. Res. Lab., Penn State Univ., University Park, PA 16804, xxd1@psu.edu) and Kenneth E. Gilbert (Univ. of Mississippi, University, MS 38677)

This paper compares a piecewise linear mapping and a cascaded conformal mapping for computing sound propagation over irregular terrain using the parabolic equation (PE) approximation. Although the piecewise linear mapping does not preserve the form of the underlying Helmholtz equation (i.e., it is not conformal), it does preserve the form of the narrow-angle PE. Further, it is shown that the correct impedance boundary condition at the ground surface can be closely approximated with the piecewise linear mapping. Compared to the cascaded conformal mapping, the piecewise mapping is extremely simple and easy to implement. It involves only a phase multiplication at the end of each range step. To test the accuracy of the piecewise linear mapping, it is compared with a cascaded conformal mapping for propagation over a single hill at various frequencies. Propagation predictions for more general terrain using the piecewise linear mapping are presented and discussed. [Work supported by the Federal Aviation Administration.]

Session 5aPAb

Physical Acoustics: Topics in Nonlinear Acoustics

David T. Blackstock, Chair

Dept. of Mechanical Engineering, Univ. of Texas at Austin, 1 University Station, Austin, TX 78712-0292

Contributed Papers

10:30

5aPAb1. Clayton H. Allen's discovery of nonlinear acoustic saturation. David T. Blackstock (Appl. Res. Labs. and Dept. Mech. Eng., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029)

In 1950 Clayton H. Allen (1918–2004) discovered nonlinear acoustic saturation. His Penn State doctoral thesis, "Finite Amplitude Distortion in a Spherically Diverging Sound Wave in Air," is about experiments on intense sound produced by a 12.2-cm diameter baffled piston vibrating at 14.6 kHz. At $r=3.8$ cm (the next-to-last nearfield peak) the maximum SPL produced was 161 dB (re $20 \mu\text{Pa}$). The farfield began at about 50 cm. Measurements were made of the fundamental and 2nd–6th harmonics out to $r=200$ cm. His received level versus source level curves at $r=200$ cm show the now familiar plateau that is reached as nonlinear effects limit the achievable SPL for a given distance at a given frequency. Clay's discovery is quite remarkable because (1) it was totally unexpected and (2) it came long before any theoretical predictions of the phenomenon were developed. Alas, however, the work was never published except in the Army Signal Corps report from Penn State, "Atmospheric Physics and Sound Propagation, Final Report for Period July 1, 1945 to May 20, 1950," and in two papers presented orally at the Penn State ASA Meeting in June 1950. As a result, Clay's remarkable work has gone unappreciated for many years. This presentation is dedicated to his memory.

10:45

5aPAb2. Nonlinear enhancement and saturation phenomena in focused ultrasound beams of various geometry. Vera A. Khokhlova, Marina S. Basova (Dept. of Acoust., Faculty of Phys., Moscow State Univ., Moscow, Russia), Michael R. Bailey, and Lawrence A. Crum (Univ. of Washington, Seattle, Washington)

The effects of nonlinear enhancement of focusing gain and saturation are studied and compared for high-intensity focused ultrasound sources with an initial Gaussian shading and uniform amplitude distribution. Simulations are performed using the Khokhlov Zabolotskaya (KZ) nonlinear parabolic equation for weakly dissipative medium in a wide range of source linear focusing gains and source pressure amplitudes, including the strongly nonlinear regime with shocks. An artificial absorption proportional to the fourth power of frequency or an asymptotic frequency-domain approach is employed in the algorithm in order to reduce the number of harmonics for accurate modeling of strongly distorted waveforms with shocks. The effect of focusing gain and amplitude shading of the source on nonlinear enhancement of acoustic energy concentration and saturation levels at the focus is discussed. It is shown that nonlinear enhancement of focusing gain is different for different values of linear gain, different spatial distributions of the source amplitude, and different parameters of acoustic field. The levels of nonlinear saturation at the focus are obtained for very high source amplitudes. The results of simulations give lower enhancement and higher saturation levels compared to the known approximate analytic predictions. [Work supported in part by NIH Fogarty and NSBRI.]

11:00

5aPAb3. Numerical simulation of nonlinear phenomena using a dispersion-relation-preserving solution of the Navier-Stokes equations. Mark S. Wochner and Anthony A. Atchley (Grad. Program in Acoust., The Penn State Univ., 217 Appl. Sci. Bldg, University Park, PA 16802)

A modified Navier-Stokes equation set that includes classical absorption and molecular relaxation is solved using a fourth-order Runge-Kutta scheme in time and a fourth-order dispersion-relation-preserving algorithm in space. The algorithm is applied to examine nonlinear phenomena such as waveform evolution of a high amplitude sine wave to old age, shock coalescence, and evolution of arbitrary time waveforms. The results are compared to analytical solutions whenever possible. The drawbacks of such an approach are discussed along with possible solutions to problems associated with the numerical method.

11:15

5aPAb4. Parabolic approximation versus geometrical acoustics for describing nonlinear acoustic waves in inhomogeneous media. Vera A. Khokhlova, Mikhail V. Averianov (Dept. of Acoust., Faculty of Phys., M. V. Lomonosov Moscow State Univ., Leninskie Gory, Moscow 119992, Russia), Robin O. Cleveland (Boston Univ., Boston, MA 02215), and Philippe Blanc-Benon (Ecole Centrale de Lyon, 69134 Ecully Cedex, France)

Propagation of intense periodic acoustic waves in inhomogeneous media is studied in the nonlinear geometrical acoustics (NGA) approximation and using nonlinear parabolic equation (NPE). Various types of 2-D inhomogeneities are considered, such as a phase screen, single Gaussian inhomogeneities, and random inhomogeneous media. Distributions of acoustic rays are obtained by numerical solution of the eikonal equation. Pressure field patterns are calculated numerically with account for nonlinear effects and diffraction using a frequency-domain algorithm for the NPE. The location of caustics and shadow zones in the ray patterns are compared with the results of the parabolic model for the areas of increased and decreased sound pressure. Both linear and nonlinear propagation is investigated in order to reveal the validity of NGA in predicting the acoustic field structure and to better understand how the combined effects of inhomogeneities, diffraction, and nonlinearity determine the overall peak and average parameters of the acoustic field. It is shown that NGA does not accurately represent all the locations of enhanced or reduced acoustic pressure even for single scattering inhomogeneities, and the discrepancies become larger for smaller size inhomogeneities and at longer distances in random inhomogeneous medium. [Work supported by CNRS, RFBR, and NIH Fogarty.]

5aPAb5. Measurements of the rate of change of the power spectral density due to nonlinearity in one-dimensional propagation. Lauren E. Falco, Kent L. Gee, and Anthony A. Atchley (Grad. Program in Acoust., Penn State Univ., 217 Appl. Sci. Bldg., University Park, PA 16802-5018)

The influence of nonlinear effects in the propagation of jet noise is typically characterized by examining the change in the power spectral density (PSD) of the noise as a function of propagation distance. The rate of change of the PSD is an indicator of the importance of nonlinearity. Morfey and Howell [AIAA J. **19**, 986–992 (1981)] introduced an analysis technique that has the potential to extract this information from a measurement at a single location. They develop an ensemble-averaged Burgers equation that relates the rate of change of the PSD with distance to the quantity Q_{p^2p} , which is the imaginary part of the cross-spectral density of the pressure and the square of the pressure. Despite its potential applicability to jet noise analysis, the physical significance and utility of Q_{p^2p} have not been thoroughly studied. This work examines Q_{p^2p} for the one-dimensional propagation of plane waves in a shock tube. The use of such a simple, controlled environment allows for a better understanding of the significance of Q_{p^2p} . [Work supported by the National Science Foundation, the Office of Naval Research, and the Strategic Environmental Research and Development Program.]

5aPAb6. Nonlinear standing waves in shaped resonators driven by a piston. Cheng Luo, Xiaoyang Huang, and Nam Trung Nguyen (School of Mech. and Production Eng., Nanyang Technolog. Univ., 50 Nanyang Ave., Singapore 639798)

The nonlinear standing waves in a shaped resonator driven by a piston are investigated analytically and experimentally. In the study, the resonator is an exponentially expanded horn with a piston oscillating at its large end. The air pressure waves inside the resonator are obtained by solving a 1-D nonlinear wave equation with the Galerkins method, taking into account the moving boundary due to the piston. The simulation results, in terms of the pressure waveforms, the pressure amplitudes, and the resonance frequency shift, show that the nonlinear standing waves generated by the piston have characteristics similar to those in the same resonator but under entirely shaking condition. The experiment is also conducted on an exponentially expanded resonator. The small end of the resonator is sealed and the big end is connected to a 5-inch 40 W loudspeaker, functioned as the driving piston. The resonance frequency shift is observed, and high amplitude pressures, up to 0.3×10^5 Pa, have been detected at the small end, which is 20 times higher than that at the big end. The experimental results agree well with the simulated values.

FRIDAY MORNING, 20 MAY 2005

REGENCY C, 8:00 A.M. TO 12:00 NOON

Session 5aPP

Psychological and Physiological Acoustics: Auditory Perception and Psychophysics II (Poster Session)

Neal F. Viemeister, Chair

Dept. of Psychology, Univ. of Minnesota, 75 East River Rd., Minneapolis, MN 55455

Contributed Papers

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon.

5aPP1. Sensitivity to stimulus distribution characteristics in auditory categorization. Sarah C. Sullivan (Dept. of Psych., Univ. of Texas, Austin, TX 78712-0187), Andrew J. Lotto (Boys Town Natl. Res. Hospital, Omaha, NE 68131), Elizabeth T. Newlin, and Randy L. Diehl (Univ. of Texas, Austin, TX 78712-0187)

Several experiments were performed to examine the ability of humans to categorize sounds as a function of training distribution characteristics. Participants were presented non-speech sounds randomly sampled from two overlapping distributions. The sounds consisted of 25 narrow-band noise bursts varying in center frequency from 1000–1360 Hz. Two conditions were created by varying the ratio of stimuli in each category (i.e., prior probabilities of each category), resulting in different ideal boundaries (maximizing accuracy). Participants were asked to categorize the sounds as either *A* or *B* and feedback was provided. In one experiment, prior probabilities were altered mid-session without alerting the subjects. Performance was tracked by plotting identification functions and noting boundary placement for each individual block. After only six blocks of training (~35 min), most subjects had established optimal or near-optimal boundaries. Identification function slopes were calculated and found to be steeper than training distribution slopes; suggesting that listeners established criterion-like boundaries as opposed to performing probability matching. The fact that listeners adjusted their responses in accordance to distribution changes within relatively short periods of time

demonstrates the perceptual systems effectiveness at optimizing categorization based on distribution characteristics. [Work supported by NSF and NIDCD.]

5aPP2. Hearing silent shapes: Identifying the shape of a sound occluding surface. Ryan L. Robart and Lawrence D. Rosenblum (Univ. of California, Riverside, Riverside, CA, 92521, rosenblu@citrus.ucr.edu)

While most psychoacoustics is concerned with perception of sound sources, there is evidence that surfaces which reflect or occlude sound can also be detected and can guide behavior [e.g., M. S. Gordon and L. D. Rosenblum, J. Acoust. Soc. Am. **107**, 2851 (2000)]. While there is also evidence that listeners can hear the shape of sound reflecting surfaces [e.g., C. E. Rice, Science **155**, 655–664 (1967)], it is not known whether the shape of sound occluding surfaces can be heard. In a series of experiments, blindfolded listeners were asked to judge the shape of surfaces (of equal area) which occluded a set of loudspeakers emitting white noise. Overall, listeners were successful at this task, with some listeners showing near perfect performance. Follow-up experiments examined the acoustic information supportive of this skill. The findings suggest a type of auditory sensitivity not often considered in the psychoacoustics literature.

5aPP3. Comparing linear regression models applied to psychophysical data. Zhongzhou Tang, Virginia M. Richards (Dept. of Psych., Univ. of Pennsylvania, Philadelphia, PA 19104, richards@psych.upenn.edu), and Andrew Shih (Univ. of Pennsylvania, Philadelphia, PA 19104)

Relative weights for a profile analysis task were obtained using four regression/classification models; correlation coefficients, linear regression, logistic regression and probit regression. The aim of the study was to examine the impact of the choice of model on the accuracy and the efficiency with which the relative weights were determined. A yes/no task was used with observers indicating whether or not there was an increment in level to the central component of an 11-component standard. On each presentation the amplitudes of the individual components of the complex were randomly perturbed using draws from a normal distribution. When a large number of trials (1250) were used to estimate the relative weights, the four methods generated nearly identical weight estimates. When smaller numbers of trials were used (112), the different methods generated patterns of relative weights that were largely similar, and the patterns deviated only modestly from the large-number solution. In terms of efficiency, the error boundaries of the different methods were nearly indistinguishable. All in all, the number of trials needed to obtain statistically significant weights is sufficiently large that there is no apparent advantage of using one method over the others. [Work supported by NIH/NIDCD.]

5aPP4. The difference in auditory memory between young and aged listeners with normal hearing. Bruce Schneider, James Qi (Dept. of Psych., Univ. of Toronto at Mississauga, Mississauga, ON, Canada L5L 1C6, bschneid@credit.erin.utoronto.ca), Juan Huang, Xihong Wu, and Liang Li (Peking Univ., Beijing 100871, China)

The perceptual fusion of a noise with its delayed copies implies that there is a memory for maintaining a detailed representation of an arbitrary waveform. This auditory memory would be important for perceptually grouping correlated sounds and segregating uncorrelated sounds in noisy, reverberant environments, in which older listeners find it much more difficult than younger listeners to process signals. To determine the temporal extent of the auditory memory and whether it is affected by aging, this study investigated the detection of a break in correlation between two correlated broadband noises in younger and older listeners with normal hearing. The results show that younger listeners could detect a 100-ms break in correlation up to interaural delays ranging from 6.3 to 23.0 ms, suggesting that higher-order central mechanisms beyond the brainstem delay lines are likely to be involved in maintaining a memory trace of the fine details of the acoustic waveform. Aged listeners, however, could detect the break up to interaural delays ranging only from 6.7 to 9.7 ms, indicating the age-related decline in auditory memory. This decline may be one of the main causes leading to perceptual difficulties experienced by older listeners in noisy, reverberant environments. [Work supported by NSERCC and MSTC.]

5aPP5. About the neglected auditory abilities within psychological tests of intelligence. Jenny Papenbrock, Kristin Seidel, Susanne Weis, and Heinz-Martin Suess (Dept. of Psych., Univ. of Magdeburg, Pfaelzer Platz, Geb. 24, Postfach 4120, 39016 Magdeburg, Germany, si-projekt@gse-w.uni-magdeburg.de)

Auditory abilities play an important role in human intellectual abilities. In his model of human cognitive abilities Carroll (1993) identified a separate domain of auditory ability. Despite their theoretical and practical relevance, these abilities were mainly excluded in psychological tests. In order to compensate for this lack we developed a test which is intended to complement existing intelligence tests. To measure general auditory abilities already existing materials (Stankov and Horn, 1980) were used along with newly developed auditory tasks. Our test includes nonverbal tasks containing only tones as well as auditory textual tasks using spoken words. We expect that the auditory tasks measure a separate ability and therefore can make an important contribution to complement already existing intel-

ligence tests. To demonstrate that the test is able to measure different aspects of auditory abilities pure tone tasks should be distinguishable from auditory textual tasks. 120 subjects worked on the auditory test as well as on a well-established test of intelligence (Test for the Berlin Intelligence Structure Model, BIS-4; Jger, S & Beauducel, 1997) which operates with written, numerical and figural material. Main results of our study are reported and implications for future research concerning psychological ability tests are discussed.

5aPP6. Contributions of individual components to the overall loudness of a multi-tone complex. Lori Leibold, Samar Khaddam, and Walt Jesteadt (555 North 30th St., Omaha, NE 68131, leiboldl@boystown.org)

The contributions of individual components to the overall loudness of a multi-tone complex were examined in a two-interval, loudness-matching task. Stimuli were five-tone complexes centered on 1000 Hz, with six different logarithmic frequency spacings, ranging from a frequency ratio of 1.012 to 1.586. Stimuli were presented for 300 ms (10 ms rise/fall). The overall level of the standard complex was fixed at 60 dB SPL (53 dB/component). Levels of the individual components of the comparison complex were selected at random from a rectangular distribution with a range of 20 dB. Ten 100-trial blocks were run for each frequency-spacing condition. Perceptual weights were computed by calculating the point-biserial correlation between the difference in the level of each component across the two intervals and the subjects response. When all components fell within the same critical band, weights varied little across components. In contrast, the range of weights increased with increasing frequency separation, with increasing weight given to the lowest and highest frequency components. The audibility of individual components was determined by measuring detection thresholds for each component in the presence of the remaining four components. A strong relation between component weight and audibility was observed for all but the narrowest spacing conditions.

5aPP7. Differences in loudness of tone complexes with positive and negative Schroeder phase. Manfred Mauermann and Volker Hohmann (Medizinische Physik, Institut für Physik, Fakultät V, C.v.O Universität Oldenburg, 26129 Oldenburg, Germany)

Tone complexes with positive ($m+$) and negative Schroeder phase ($m-$) have an identical long-term spectrum, the same temporal envelope but are inverted in time. They show large differences in masking efficiency most probably being related to a different cochlear representation. The current study investigates to which extent loudness perception is affected similarly by the different phase characteristic of $m+/m-$ stimuli. Therefore, the loudness of $m+/m-$ stimuli (1.6 octave bandwidth) was matched in seven normal hearing and three hearing impaired subjects. In the first experiment the fundamental frequency f_0 of the tone complexes was varied from 2–1536 Hz for different center frequencies at a soft level of the reference signal ($m-$). In normal hearing subjects the $m+$ stimuli need a 6 dB higher level to be perceived as equally loud as the respective $m-$ stimuli (for f_0 in the range of 24–96 Hz). In the second experiment the difference in loudness of $m+/m-$ was investigated as function of the of the reference-stimulus level (5–80 dB SL) at a center frequency of 2660 Hz and f_0 at 48 Hz. The largest differences in loudness were found for levels between 40–60 dB, clearly reduced for higher and lower levels of the reference stimulus.

5aPP8. Perception of combined intensity and frequency contours by normal-hearing and hearing-impaired listeners. Marjorie R. Leek, Michelle R. Molis (Army Audiol. & Speech Ctr., Walter Reed Army Medical Ctr., 6900 Georgia Ave, N.W., Washington, DC 20307-5001, Marjorie.Leek@naamedd.army.mil), and Jennifer J. Lentz (Indiana Univ., Bloomington, IN 47408)

When a speaker utters a consonant-vowel or vowel-consonant syllable, the frequency transitions are critical to identification of the sound. However, an intensity contour is also superimposed on the frequency glides, as the mouth opens or closes. In this study, the interaction of frequency and intensity contours was explored using a frequency glide discrimination task. Offset frequency difference limens (offset DLs) were measured for upward and downward frequency glides in two frequency regions, and with three intensity contours, increasing from silence, decreasing to silence, and steady-state. For both normal-hearing (NH) and hearing-impaired (HI) subjects, larger offset DLs were observed for the high frequency stimuli than for the lower frequencies, and for upward-gliding stimuli than for downward frequency glides. Amplitude contour had little effect on the NH data, but did influence offset DLs for HI subjects. These findings indicate that for these stimuli, the interaction of frequency and amplitude contours plays only a small role in the perception of transition-like glides for NH listeners, but may affect the perception of frequency transitions by HI listeners. [Work supported by NIDCD.]

5aPP9. Processing two stimuli simultaneously: Switching or sharing? Frederick J. Gallun, Christine R. Mason, and Gerald Kidd, Jr. (Hearing Res. Ctr. and Comm. Disord., Boston Univ., 635 Commonwealth Ave., Boston, MA 02215)

This study explored whether listeners can perform two auditory tasks simultaneously or whether they must switch rapidly between tasks. The stimuli were two simultaneous speech streams (sentences) processed into sets of mutually-exclusive narrow frequency bands. One was presented to each ear and listeners reported either the keywords from the sentences (Task 1, recognition) or simply whether or not a sentence had been presented (Task 2, detection). For both sentences and tasks, overlapping bands of noise were presented at a level that reduced performance to below 90% correct in a single-stream reference condition. Listeners were informed of which stream to report either before or after stimulus presentation. The two tasks were either the same (Task 1 at both ears) or different (Task 1 on the right, Task 2 on the left). The effect of having the opportunity to rapidly switch between sentences was examined by presenting full sentences or only the keywords. Interference occurred primarily when the tasks were the same in both ears and was greatest when the target ear was not specified in advance. Presenting only the keywords hurt single-stream performance but did not increase interference in the different-task case, arguing against a rapid-switching explanation. [Work supported by NIH/NIDCD.]

5aPP10. Explaining two-tone suppression and forward masking data using a compressive gammachirp auditory filterbank. Toshio Irino (Faculty of Systems Eng., Wakayama Univ., 930 Sakaedani, Wakayama 640-8510, Japan) and Roy Patterson (Univ. of Cambridge, Cambridge, CB2 3EG, UK)

The gammatone filter was imported from auditory physiology to provide a time-domain version of the roex auditory filter and enable the development of a realistic auditory filterbank for models of auditory perception. The gammachirp auditory filter was developed to extend the domain of the gammatone auditory filter and simulate the changes in filter shape that occur with changes in stimulus level. Recently, the gammachirp was extended to explain the level-independent frequency glide of the impulse response and a large volume of simultaneous masking data quantitatively. Although this could be implemented with a static filter, we used a time-varying filter whose active component is an IIR asymmetric compensation filter. In this case, it is necessary to estimate the signal level that controls the level dependency, and explain how the level measurement

produces the time course of threshold observed in forward masking experiments. In this talk, we propose a new implementation of the time-varying, compressive gammachirp auditory filterbank to explain two-tone suppression and forward masking data qualitatively. We will also show that the system can resynthesize compressed speech sounds and so provide the basis for a gammachirp analysis/synthesis filterbank. [Work supported by GASR(B)(2) No. 15300061, JSPS.]

5aPP11. Behavioral estimates of level-dependent shifts in the vibration pattern of the apical region of the basilar membrane. Luis F. Barrios, Enrique A. Lopez-Poveda, and Ana Alves-Pinto (Instituto de Neurociencias de Castilla y Leon, Univ. of Salamanca, Avda. Alfonso X El Sabio s/n, 37007 Salamanca, Spain)

The best frequency of any given site on the BM shifts to a different value as the stimulus level increases. For basal sites, the shift occurs towards lower frequencies. For apical sites, however, the direction of the shift is unclear. Auditory nerve data [e.g., Rose *et al.*, *J. Neurophysiol.* **34**, 685–699 (1971)] suggest that shifts may not occur or may occur toward higher frequencies. The present work investigates which is the case in human. To do it, psychophysical tuning curves were measured for three normal-hearing listeners using forward masking. The level of a masker tone required to just mask a fixed, low-level probe tone was measured for different masker-probe time intervals. The duration of the interval was adjusted as necessary to obtain PTCs for the widest possible range of masker levels. Probe frequencies of 125, 250, and 500 Hz were considered. Masker frequencies ranged from 0.5 to 2.0 times the probe frequency. A clear shift toward lower frequencies occurred for all probe frequencies for one listener only. For the other two listeners, shifts were not observed for any probe frequency. No clear shifts toward higher frequencies were observed. [Work supported by Spanish FIS PI020343 and G03/203.]

5aPP12. Modeling the influence of the cochlear nonlinearity on estimates of psychophysical tuning. Xuedong Zhang and Andrew J. Oxenham (Res. Lab. of Electron., MIT, Cambridge, MA 02139)

Most behavioral measures of human frequency selectivity have been made with simultaneous masking and the notched-noise technique. The resulting filter shapes may be influenced by the effects of cochlear nonlinearity, such as suppression. Forward masking may provide a measure that is more comparable to neural tuning curves, because it does not involve stimuli that interact with each other along the basilar membrane. This study investigated the extent to which cochlear nonlinearities can account for differences in results between forward and simultaneous masking. The model was constructed using a nonlinear auditory filter, a sliding temporal integrator, a logarithmic transform and a template mechanism. The effects of compression and suppression on psychophysical performance were simulated by varying the relevant parameters of the model auditory filter. The psychophysical results were simulated for both forward and simultaneous masking, using the same parameters and tracking procedure as in the behavioral studies. The results provide a detailed evaluation of the role of compression and suppression in the models predictions of psychophysical tuning and assist in the development of the refined nonlinear cochlear models for human. [Work supported by the ASA Hunt Fellowship and NIH R01DC03909.]

5aPP13. Estimating cochlear-filter shapes, temporal-window width and compression from tone-sweep detection in spectral and temporal noise gaps. Gaston Hilkhuisen, Tammo Houtgast, and Johannes Lyzenga (Dept. of Otolaryngol., VU Univ. Medical Ctr., de Boelelaan 1117, 1081 HV Amsterdam, The Netherlands, G.Hilkhuisen@vumc.nl)

A test, designed for naive listeners, measured tone-sweep detection in noise with either spectral or temporal gaps. For normal-hearing (NH) listeners, detections in spectral gaps depended on level, which can be ex-

plained from Outer-Hair-Cell (OHC) activity. At low levels, OHC activity increased frequency-selectivity by amplifying the signal in the spectral gap, improving the signal-to-noise ratio (SNR). Relative to the broad passive cochlear filter, OHC activity decreased with rising signal levels. In consequence, SNRs decreased and detection deteriorated. Similarly, decreasing OHC activity may explain the observed level dependence of detection thresholds in temporal gaps. At low and high intensities, signal and noise were equally amplified. Detection was best at intermediate levels when the low-level signal in the temporal gap was amplified more than the high-level noise. All effects are modeled using a one-parameter time window with decaying-exponential shape preceded by a simplified dual-resonance non-linear (DRNL) filter. The filter contains two parallel, one-parameter, Rounded Exponential filters: a broad filter representing passive cochlear filtering and a narrow one, combined with a level-dependent amplifier, representing OHC activity. From estimated filter and time-window widths and OHC amplification, compression curves are derived. Additionally, results for hearing-impaired listeners will be presented.

5aPP14. Speech understanding in noise: Contributions of compressive nonlinearities in the basilar-membrane response. Amy R. Horwitz, Jayne B. Ahlstrom, and Judy R. Dubno (Dept. of Otolaryngol.-HNS, Medical Univ. of South Carolina, 135 Rutledge Ave., P.O. Box 250550, Charleston, SC 29425, horwitar@musc.edu)

The contribution of compressive nonlinearities in the basilar-membrane response was assessed by measuring speech recognition in noise as a function of noise level and growth of masking for tones. Consonant recognition was measured in interrupted noise at overall masker levels of 47–77 dB SPL. Additionally, thresholds for a 10-ms, 2.0-kHz tone were measured in a masker ranging in level from 40 to 85 dB SPL and centered at 1 kHz. Subjects were younger and older adults with normal hearing. With speech at a constant level and the masker level varying, it was hypothesized that the response of the basilar membrane to the masker would be linear at lower levels and compressed at medium to higher levels, resulting in less effective masking at higher masker levels. Consistent with the hypothesis, differences between observed and predicted consonant-recognition scores, determined using AI calculations, varied as a function of masker level, especially for younger subjects. At lower masker levels, observed scores declined as predicted as masker level increased; with further increases in masker level, scores declined less than predicted. For tonal growth-of-masking functions, breakpoints were correlated with quiet thresholds: both were significantly higher for older than younger subjects. [Work supported by NIH/NIDCD.]

5aPP15. Psychometric-function slopes for forward-masked tones in listeners with cochlear hearing loss. Kim S. Schairer, Jessica Messersmith, and Walt Jesteadt (Boys Town Natl. Res. Hospital, 555 North 30th St., Omaha, NE 68131)

Slopes of psychometric functions (PF) for forward-masked tones in normal-hearing (NH) listeners decrease as a function of signal threshold in both on- and off-frequency masker conditions. This result is attributed to nonlinear cochlear response growth, and has been used to demonstrate comparable compression at 4000 and 250 Hz. The current study further tested the hypothesis that the slope effect is due to nonlinearity. In hearing-impaired (HI) listeners who presumably lack cochlear nonlinearity, PF slopes should remain steep across signal threshold levels. Four NH and six HI listeners with a range of hearing thresholds were tested in on- and off-frequency masker conditions, with signals of 4000 and 250 Hz. Otoacoustic emissions were measured to verify absent or decreased cochlear nonlinearity in the HI listeners. In general, PF slopes were steeper for the HI than NH listeners. One subject had normal hearing at 250 Hz, and hearing loss at 4000 Hz. Her PF slopes were comparable to the mean slopes across NH subjects in the 250-Hz signal conditions, but were elevated in comparison to the NH subjects in the 4000-Hz conditions. These results are consistent with predictions from a model of forward masking that incorporates cochlear nonlinearity as one of its stages.

5aPP16. Masker variability effects in backward-masked frequency discrimination. Blas Espinoza-Varas and Hyunsook Jang (Commun. Sci. and Disord., Univ. Oklahoma Health Sci. Ctr., Oklahoma City, OK 73190)

Masker variability effects have been studied mostly in detection tasks with simultaneous maskers; this paper reports variability effects in a frequency discrimination threshold (FDT) paradigm with backward maskers. FDTs for sinusoidal targets were measured unmasked (or in isolation) and in two backward-masked conditions: one included within-trial masker frequency variability, the other did not. The 1500-Hz, 40- or 80-ms targets, and 500-, 1500-, or 2500-Hz, 80-ms maskers were presented at 70-dB SPL with a 20-ms ISI. A standard and two comparisons were displayed, respectively, in the 1st, 2nd, and 3rd observation interval of a 3I/2AFC task. One comparison, chosen randomly with 0.5 probability, contained an increment in target frequency as specified by an adaptive rule that estimated 71 percent-correct FDTs. Identical increments in masker frequency were added to both comparisons in within-trial variability conditions, but not in conditions without variability. In all conditions, the task was to determine which comparison contained the target frequency increment. Masked no-variability FDTs were higher than unmasked FDTs, but training nullified this effect. Masker variability induced greater FDT elevations with 40- than with 80-ms targets, and in conditions having the same rather than different target-masker frequencies. Prolonged training nullified all masker variability effects.

5aPP17. Hearing relative phases for two harmonic components. D. Timothy Ives, Roy D. Patterson (C.N.B.H., Physio. Dept., Univ. of Cambridge, Downing St., Cambridge, CB2 3EG, UK, tim.ives@mrc-cbu.cam.ac.uk), and H. Martin Reimann (Univ. of Berne, 3012 Berne, Switzerland)

There is a new wavelet model of basilar membrane motion which predicts that existing roex and gammatone filterbanks underestimate the interaction of harmonically related components in complex tones. This interaction appears as a modulation of the basilar membrane motion associated with a higher harmonic by the presence of a lower harmonic; the period of the modulation is that of the difference frequency, or the fundamental of the implied harmonic series. The modulation depth is largest for stimuli whose spectra have a similar energy distribution as natural sounds: a high frequency roll-off of about 12-dB per octave. The strength of the modulation frequency is dependent on the relative phase of the components, which predicts that listeners will be able to hear a relative phase change. A series of experiments were undertaken to determine the threshold for a change in the relative phase of the components. The results of the experiments show that a change in the relative phase is detectable for components that are generally believed to be resolved on the basilar membrane and therefore should not interact. [Work supported by the U.K. Medical Research Council (G9901257).]

5aPP18. The fading of auditory memory. Liang Li, Juan Huang, Lingzhi Kong, Ying Huang (Dept. of Psych., Speech and Hearing Res. Ctr., Peking Univ., Beijing 100871, China liangli@pku.edu.cn), Xihong Wu, Jing Chen, Qiang Huang, Yuan Yao (Peking Univ., Beijing 100871, China), James Qi, and Bruce Schneider (Univ. of Toronto at Mississauga, Mississauga, ON, Canada)

Due to auditory memory, the auditory system is capable of maintaining a detailed representation of arbitrary waveforms for a period of time, so that a broadband noise and its delayed copies can be perceptually fused. This auditory memory would be critical for perceptually grouping correlated sounds and segregating uncorrelated sounds in noisy, reverberant environments. Its fading process over time was investigated in the present study at the behavioral level, using a break in correlation (BIC, a drop of inter-sound correlation from 1.00 to 0 and then return to 1.00) between two correlated broadband noises. The results show that with the rise of inter-sound delay from 2 to 10 ms under either headphone-stimulation or loudspeaker-stimulation conditions, the shortest BIC duration necessary

for listeners to correctly detect the occurrence of the BIC increased rapidly. This elevation in the duration threshold was faster under the headphone-stimulation condition than the loudspeaker-stimulation condition. Also, the listeners reaction time in response to the BIC but not that to a comparable silent gap elongated quickly with the increase in the inter-sound delay from 1 to 8 ms. Thus the auditory memory of fine structures fades rapidly after the sound waves are received. [Work supported by MSTC and NSERCC.]

5aPP19. Evidence for a cancellation mechanism in perceptual segregation by differences in fundamental frequency. John F. Culling, Gareth M. Linsmith, and Tracy L. Caller (School of Psych., Cardiff Univ., Tower Bldg., Park Pl. Cardiff, CF10 3AT, UK)

Identification of target speech material is better when it differs in fundamental frequency (F_0) from interfering speech material. Two experiments tested the idea that this effect is mediated by a cancellation process that perceptually removes sound on one F_0 . Expt. 1 used monotonized target speech at 112.5 Hz F_0 against two monotonized interferers. The F_0 s of the interferers were either both 112.5 Hz, both 126 Hz, or 100 and 126 Hz. Speech reception thresholds (SRTs) were about 3.5 dB lower when the two interferers shared a 126 Hz F_0 than in the other two conditions, which differed by only 0.3 dB. These results are consistent with a cancellation mechanism that is able to remove sounds at only a single F_0 . Expt. 2 replicated and extended Expt. 1, testing whether the results were affected by the allocation of the target and interferers to particular F_0 s. When all sentences shared a common F_0 , SRTs were now 1–1.5 dB higher than when they were all different. However, all 4 SRTs for interferers sharing the same F_0 (differing from that of the target by 2 or 4 semitones) were lower than all three of the SRTs for interferers with independent F_0 s.

5aPP20. Effects of high-frequency amplification and training on impaired-listeners' ability to discriminate redundant cues. Blas Espinoza-Varas, Shelagh Bowman-Edmundson, and Hyunsook Jang (Commun. Sci. and Disord., Univ. Oklahoma Health Sci. Ctr., Oklahoma City, OK 73190)

This paper examines effects of high-frequency amplification and training on the ability to discriminate redundant duration increments added simultaneously to the low (L) and high (H) frequency component of two-component complexes. The frequency of L (500 or 1000 Hz, 80 ms) and H (2000 or 3127 Hz, 60 ms) was chosen so as to stimulate simultaneously normal- and impaired-sensitivity regions of high-frequency sensorineural patients. The 3I/2AFC task displayed a standard followed by two comparisons. Duration increments in L , H , or both (tL , tH , or tLH), occurred randomly in comparison 1 or 2, and listeners had to decide which had an increment. The training stages were: (1) no H amplification, with $L=35$, $H=-5$ dB SL and tL , tH , and tLH ; (2) H amplification, same as before but with $H=35$ dB SL; (3) H amplification and tH discrimination training, same as stage 2, but with tH only and H in isolation or together with L ; and (4) retest of stage 2, post tH discrimination training. Only in stage 4 did redundancy yield a significant improvement in overall discrimination accuracy; amplification alone yielded little improvement.

5aPP21. Wideband reflectance in normal and otosclerotic ears. Navid Shahnaz and Karin Bork (Univ. of British Columbia, School of Audiol. & Speech Sci., 5804 Fairview Ave., Vancouver, BC, Canada V6T 1Z3, nshahnaz@audiospeech.ubc.ca)

For years immittance has been used in order to help diagnose middle ear pathologies. Specifically, multi-frequency tympanometry (MFT) is able to relate information regarding the mechano-acoustical characteristics of the middle ear system. In the past two decades a new method of middle ear measurement, wideband reflectance (WBR), has been introduced. WBR is the ratio of energy reflected from the surfaces of the ear canal and

middle ear on its way to the cochlea in relation to the energy that reaches the surface, or incident energy. This ratio is known as energy reflectance. This paper adds to the limited normative data available, as well as explores whether these normative data have a clinical utility in the diagnosis of otosclerosis. Descriptive statistics were gathered from 62 (115 ears) Caucasian normal hearing adults as well as in seven patients (seven ears) with otosclerosis. All of the otosclerotic patients in this study deviated from the normative values on at least one of the four WBR parameters of power absorption, admittance, susceptance, or conductance even when their MFT results were within normal limits. Although only seven patients were tested, these results provided evidence in favor of the utility of WBR for diagnosis of otosclerosis.

5aPP22. Sensitivity to combined frequency and amplitude speech cues by cochlear implant users. Eric W. Healy, Cheryl F. Rogers, and Allen A. Montgomery (Dept. of Commun. Sci. and Disord., Arnold School of Public Health, Univ. of South Carolina, Columbia, SC 29208, ewh@sc.edu)

Although a considerable amount of work has been directed toward examining the sensitivity of the cochlear implant (CI) user to various acoustic cues, less has examined the ability to combine cues. The sensitivity to concurrent cues was examined using an adaptive 3I-2AFC paradigm. Stimuli were synthesized three-syllable words having middle syllables that possessed an increase in fundamental frequency, amplitude or both. Sensitivity to increments in the individual cues was first measured for five subjects with a CI and five age-matched controls. These DLs were then used to equate sensitivity to the two cues and create stimuli having concurrent increments. It was found that the presence of the two cues reduced the mean DL to half that of either cue alone, for both groups of subjects. Thus, this combination of cues is similar across groups and simply additive. Individual differences in the ability of CI users to combine cues were more strongly predictive of performance on speech recognition tests than was sensitivity to either cue alone. These results have implications for the perception of the multiplicity of speech cues. [Work supported by NIDCD.]

5aPP23. Recognition of vowels that have been spectrally warped according to the frequency map for the spiral ganglion. Peter S. Popolo and Christopher W. Turner (Dept. of Speech Pathol. and Audiol., 119 WJSHC, The Univ. of Iowa, Iowa City, IA 52242, peter-popolo@uiowa.edu)

The purpose of this study was to explore the effect of a specific spectral warping on vowel recognition in cochlear implant users. The warping was designed to simulate the misalignment of frequency-to-place stimulation which occurs if the spiral ganglion cells are electrically stimulated instead of the eighth-nerve dendrites in the organ of Corti. Simulated CIS-processed vowel stimuli were generated in which the analysis and carrier band center frequencies were shifted according to the projection of the spiral ganglion cells onto the organ of Corti, derived from published data [A. Kawano, H. L. Seldon, and G. C. Clark, Ann. Otol. Rhinol. Laryngol. **105**, 701–709 (1996)]. Similar spectral warping was applied to unprocessed vowel stimuli. Listening experiments were conducted in which normal hearing subjects were presented with four types of stimuli: unprocessed-unwarped, processed-unwarped, unprocessed-warped, and processed-warped. For each condition, performance was allowed to stabilize over repeated trials with feedback, to allow for learning. Vowel recognition scores averaged across listeners were drastically reduced for the warped stimuli, regardless of whether the stimuli were CIS-processed or not. The combination of spectral warping and reduced spectral resolution resulted in the poorest intelligibility scores. These results may have implications for the design of frequency-to-place maps in cochlear implants.

5aPP24. Utilizing different channels for multiple inputs in cochlear implant processing. Bom Jun Kwon (Cochlear Americas, 400 Inverness Pkwy, Ste. 400, Englewood, CO 80112, bjkwon@gmail.com)

While cochlear implants successfully provide auditory sensation for deaf people, speech understanding through the device is compromised when there is a background noise or competing sounds, partly due to implant users' reduced ability in auditory grouping. The present study investigates whether providing multiple streams of input on different channels would facilitate auditory grouping, thereby assisting speech understanding in competing sounds. In acoustic hearing, presenting two streams of input (such as speech and noise) in spectrally separate channels gener-

ally facilitates grouping; however, in electric hearing it is difficult to predict and separating them could lead to a negative result, because channel interactions inferred from the excitation patterns are severe and the actual amount of electric current for the noise delivered to the cochlea would be much higher for a given SNR, therefore contaminating the target more effectively. Results from consonant identification measured in a variety of speech/noise conditions (same/different channels) indicate that speech understanding generally improves with separate channels, implying that implant users appear to extract speech information on the basis of spatial (channel) separation, easily overcoming the distracter from the adjacent channels with higher intensity. This also proposes a new measure of channel interactions based on auditory grouping.

FRIDAY MORNING, 20 MAY 2005

PLAZA B, 9:00 A.M. TO 12:00 NOON

Session 5aSA

Structural Acoustics and Vibration: General Vibration: Modeling, Propagation, Dissipation and Control

Linda P. Franzoni, Chair

Dept. of Mechanical Engineering and Material Science, Duke Univ., Durham, NC 27708-0300

Contributed Papers

9:00

5aSA1. Time reversal in heterogeneous flexural beams. Dany Francoeur and Alain Berry (Dept. Gen. Mecanique, Universite de Sherbrooke, 2500 boul. de l'Universite, Sherbrooke, QC, Canada J1K 2R1, Alain.Berry@USherbrooke.ca)

Time reversal of acoustic and structure-borne waves has been explored in recent years mostly for ultrasound and for nondispersive propagation, that is under frequency-independent wave velocity. In contrast, the case of time reversal in flexural beams presented here involves dispersive propagation, and is carried for frequencies below 5 kHz. The study has been started with analytical time reversal simulations in infinite homogeneous or heterogeneous beams (comprising point-mass scatterers). Experiments have also been realized on a 5 m beam with anechoic terminations and under transverse impact excitation. The time-reversal mirror was made of several thin piezoceramic elements bonded on the beam to sense the impulse response of the structure and re-emit its time-reversed image. The experimental results are in good agreement with the analytical results, and show that time spreading due to dispersive propagation of bending waves is successfully compensated by the time reversal operation. The presentation will illustrate the main results of the simulations and a comparison with the experiments.

9:15

5aSA2. Medium frequency vibration modeling of cracked plates using hierarchical trigonometric functions. Jérôme Pinonnault, Patrice Masson, Philippe Micheau (GAUS, Mech. Eng. Dept., Université de Sherbrooke, Sherbrooke, QC, Canada J1K 2R1, Patrice.Masson@USherbrooke.ca), and Nezih Mrad (Defence R&D Canada, Dept. of Natl. Defence, Ottawa, ON, Canada K1A 0K2)

A modeling tool is proposed to describe the vibration behavior of pristine and damaged plates in the medium frequency range (below 10 kHz). This tool is intended to provide a platform for the development and assessment of damage detection algorithms for aircraft structural health monitoring applications. The proposed analytical model employs a Hierarchical Trigonometric Function Set (HTFS) to characterize homogeneous plates with through cracks. This approach takes advantage of the very high order of stability of the HTFS [O. Beslin and J. Nicolas, *J. Sound Vib.* **202**, 633–655 (1997)] to approximate the effects of a small crack in a plate for all combinations of classical boundary conditions (e.g., CFSC,

CCFF, FSFS). The model is first presented and then assessed for healthy and cracked CCCC plates with eigenvalues and eigenmodes presented in the literature. For a healthy square plate, numerical results provide good agreement up to the 1000th mode while, for a cracked rectangular plate, good agreement is obtained up to the 3rd mode, corresponding to the highest mode order available in the literature. Wave propagation simulation obtained from HTFS shows the scattering around the cracks in the plates. Experimental validation of the model is conducted both in frequency and time domains for healthy and cracked plates. [Work supported by the Consortium for Research and Innovation in Aerospace in Quebec (CRIAQ) and Defence R&D Canada.]

9:30

5aSA3. An exact analytical expression of the shear coefficient in the Mindlin plate equation. Andrew Hull (Code 8212, Naval Undersea Warfare Ctr., Newport, RI 02841, hullaj@npt.nuwc.navy.mil)

This paper derives an exact analytical expression of the shear coefficient in the Mindlin plate equation for a plate of infinite extent. The equation of motion derived from the Mindlin plate equation is set equal to the equation of motion derived from the thick plate equation, and the result is a closed-form expression of the shear coefficient at all wave numbers and frequencies. A numerical example is included to illustrate the variation of the shear coefficient. It is shown that the shear coefficient is extremely dependent on wave number and only slightly dependent on frequency. Shear coefficients derived in other work are compared favorably to the values calculated by this new method at the plate flexural wave response. [Work funded by the Office of Naval Research.]

9:45

5aSA4. Structural acoustic control of plates with variable boundary conditions. Robert L. Clark, Joseph D. Sprofera (Dept. of Mech. Eng. and Mater. Sci., Duke Univ., Box 90300, Durham, NC 27708-0300), Gary P. Gibbs, and Ran H. Cabell (NASA Langley Res. Ctr., Structural Acoust. Branch)

A method for optimizing a structural acoustic control system with respect to potential variations in plate boundary conditions is provided. The assumed modes method is used to build a plate model with varying levels of rotational boundary stiffness to span a range of possible boundary conditions which can be used to capture uncertainty in the model. A trans-

ducer placement scoring process, involving Hankel singular values (HSVs), is combined with a genetic optimization routine to find spatial locations robust to boundary condition variation. Predicted frequency response characteristics are examined, and optimized results are discussed in relation to boundary condition variations. Results indicate it is possible to minimize the impact of uncertain boundary conditions in active structural acoustic control by optimizing the placement of transducers with respect to uncertainties. Both analytical and experimental results will be discussed.

10:00–10:15 Break

10:15

5aSA5. Optimal energy dissipation in a semi-active friction device. Paulin Buaka, Philippe Micheau, and Patrice Masson (GAUS, Mech. Eng. Dept., Université de Sherbrooke, Sherbrooke, QC, Canada J1K 2R1, Patrice.Masson@USherbrooke.ca)

A semi-active device is presented for vibration control using energy dissipation by dry friction at contact surfaces. Semi-active behavior is provided by two piezoelectric stack actuators driven in real time to apply a normal force on a mobile component through two friction pads. Theoretical and experimental results show that there is an optimal constant normal force to maximize the energy dissipated for the case of a harmonic disturbance. In order to improve the energy dissipation by real time control of the normal force, two nonlinear controllers are proposed: (1) the Lyapunov method leading to a nonlinear bang-bang controller law and (2) the feedback linearization approach leading to equivalent viscous friction. The implementation of both strategies is presented and both are experimentally assessed using a clamped-free beam with the semi-active device attached to the beam. It is shown that a proper choice for the parameters of the controllers leads to an increased energy dissipation with respect to the case where the normal force is constant. This dissipation is further increased by adjusting a phase shift in the nonlinear feedback loop in order to avoid a stick-slip motion of the mobile component.

10:30

5aSA6. Energy sinks: Vibration absorption by an optimal set of undamped oscillators. Ilker Koç (Mech. Eng. Dept., Carnegie Mellon Univ., Pittsburgh, PA 15213), Antonio Carcaterra (Universita Degli Studi di Roma "La Sapienza," 00184 Roma, Italy), Zhaoshun Xu, and Adnan Akay (Carnegie Mellon Univ., Pittsburgh, PA 15213)

This presentation offers the concept of energy sinks as an alternative to conventional methods of vibration absorption and damping. A prototypical energy sink envisioned here consists of a set of oscillators attached to, or an integral part of, a vibrating structure. The oscillators that make up an energy sink absorb vibratory energy from a structure and retain it in their phase-space. In principle, energy sinks do not dissipate vibratory energy as heat in the classical sense. The absorbed energy remains in an energy sink permanently so that the flow of energy from the primary structure appears to it as damping. This paper demonstrates that a set of linear oscillators can collectively absorb and retain vibratory energy with near irreversibility when they have a particular distribution of natural frequencies. The approach to obtain such a frequency response is based on an optimization that minimizes the energy retained by the structure as a function of frequency distribution of the oscillators in the set.

10:45

5aSA7. Experiments on vibration absorption using energy sinks. Adnan Akay, Zhaoshun Xu (Mech. Eng. Dept., Carnegie Mellon Univ., Pittsburgh, PA 15213), Antonio Carcaterra (Universita Degli Studi di Roma "La Sapienza," 00184 Roma, Italy), and Ilker Koç (Carnegie Mellon Univ., Pittsburgh, PA 15213)

This presentation describes experiments that demonstrate the concept of energy sinks where a set of multiple undamped linear oscillators attached to a vibrating structure can absorb most of its energy. In principle,

energy sinks do not require presence of damping in the classical sense. A set of undamped oscillators that make up an energy sink collectively absorb the vibratory energy and retain it in their phase space. Earlier optimization studies by the authors have shown the feasibility of vibration absorption and retention by energy sinks if the set of oscillators have a particular frequency distribution. Experimental results support the concept of energy sinks. Different physical realizations of energy sinks demonstrate the significance of frequency distributions and the ability of energy sinks to reduce vibration amplitude of a primary structure to which they are attached.

11:00

5aSA8. Experimental study on passive/active hybrid isolation. Zi Jun Zhang, Woo Suk Chang, Koon Meng Nyang, and Yew Wing Chan (DSO Natl. Labs., Singapore, No 20, Sci. Park Dr., Singapore 118230)

This paper discusses the control of a high stroke low stiffness nonlinear actuator which formed the key element of a smart engine mount system that provides strong support for the engine while at the same time absorbing the engine vibration energy. The actuator is made of a stacked PZT embedded in an elliptical shaped metal frame, in the horizontal direction. Due to the geometry of the frame, the displacement generated by the PZT stacks is amplified up to about 5 times in the vertical direction. However, due to the geometry of the actuator, the vertical and horizontal motions are non-linear and excite motions at multiples of the driving frequency. Feedback controllers cannot be used since the high frequency motion causes spillover problem and control becomes unstable. The filtered-X adaptive controller with sufficient high sampling rate is tested and found to be a useful and reliable controller for the actuator for suppressing the vibration to almost ambient noise level. This actuator requires relatively lower driving voltage level and delivers larger stroke range. This would be more suitable for vibration isolation application where the actuator is serially connected with passive dampers.

11:15

5aSA9. Comparison between control strategies for active constrained layer damping treatment to control the sound radiation from a vibrating structure. Hui Zheng, Chun Lu, and Heow Pueh Lee (Inst. of High Performance Computing, 1 Sci. Park Rd., #01-01 The Capricorn, Sci. Park II, Singapore 117528)

A comparative study is presented of three control strategies for active constrained layer damping (ACLD) treatments to control the sound radiation from a vibrating planar structure. The first control strategy is one commonly used in most existing studies of ACLD for structural vibration control; i.e., the error signal to be minimized for the controller is the vibration response sensed by point transducers, and a proportional derivative controller is employed where the sensor signal and the voltage output is related by real-valued feedback gains. The second control strategy is similar to the first, except that the real-valued control gains are substituted by complex-valued ones. In the third control strategy, the discrete structural acoustic sensing approach is introduced for observing the sound radiation from the structure, and the estimated sound power constitutes the controller input. The control gains aiming to reduce the sound radiation from a simply-supported beam are optimized respectively for the three control strategies, and the control performances are compared. Numerical results show that using the complex-valued control gain in the controller design is always better than using real-valued gain. Furthermore, the ACLD treatments adopting the third control strategy require lowest control efforts.

5aSA10. Vibration control of optomechanical components. Vyacheslav M. Ryaboy (Newport Corp., 1791 Deere Ave., Irvine, CA, 92606, vyacheslav.ryaboy@newport.com)

The purpose of optomechanical components is to anchor optical elements (such as light sources, lenses, mirrors, etc.) in space so that the optical paths would be unperturbed by environmental impacts, first of all by vibration. Dynamic alignment of components is a main design optimization criterion for the optomechanical systems. An optical table or breadboard is often used as a common base for the whole optomechanical assembly. It can be sufficiently isolated from the floor vibration, but its own flexural resonance vibrations may still cause misalignments. The paper estimates various ways of vibration damping of optical tables, including new methods of passive damping combined with motion transformation, as well as active damping. Optical rods and posts offset optical elements from the table. These components have their own resonance properties. The paper presents analytical estimates and optimization methods for introducing damping materials in the optical posts. Experimental results comparing dynamic properties of damped and non-damped structures will be presented.

5aSA11. Whole body vibration measurements on forklift trucks. Alberto Behar (Noise Control, 45 Meadowcliffe Dr, Scarborough, ON, Canada, M1M 2X8 behar@sympatico.ca) and Steven Libich (WESA)

Data on acceleration values on forklift trucks related to Whole Body Vibration are notably missing in the literature. Although there are several standards that provide guidance on how measurements have to be performed, they are somehow conflicting and difficult to implement since they require simultaneous measurement and processing of data in the three axes, x , y and z . Standards also provide limit values for safety of the personnel exposed to the vibrations. The development of new instruments has made these kind of measurements much easier to perform and to interpret the results, since they include filters following the standards and allow for the use of tri-axial accelerometers. This paper describes measurements done on 11 forklift trucks in a real-life situation, while performing tasks normal for their use. The trucks are of the standing-operator type. The accelerometers were mounted on the body of the truck, so that vibrations were representative for what the operators were exposed to. The three-axes orientation of the accelerometer were taken into account. The paper will present results of the measurements and their assessment following the existing ISO and BS Standards.

FRIDAY MORNING, 20 MAY 2005

REGENCY D, 8:00 A.M. TO 12:00 NOON

Session 5aSC

Speech Communication: Intelligibility and Studies of Populations with Speech and Hearing Disorders (Poster Session)

Sigfrid D. Soli, Chair

House Ear Inst., 2100 W. Third St., Los Angeles, CA 90057

Contributed Papers

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon.

5aSC1. Evaluation of effect of presbycusis on speech intelligibility by several kinds of speech test in rooms. Hiroshi Sato (Inst. for Human Sci. & Biomed. Eng., Natl. Inst. of Adv. Industrial Sci. and Technol., 1-1-1 Higashi, Tsukuba, Japan) and Hayato Sato (Kobe Univ., Rokko, Nad, Kobe 657-8501, Japan)

Word recognition tests with logatom and word familiarity controlled word lists and sentence intelligibility test in simulated sound fields with noise and/or reverberation were carried out to assess the effect of hearing loss due to aging on speech communication in rooms. The result demonstrates that (1) speech recognition scores of elderly listeners are 25% lower than those of young adults for any kinds of speech test. This difference is equal to the 5 dB increase of ambient noise for elderly listeners. (2) Detailed speech recognition process in noise and/or reverberation is described with scores of various kinds of speech test. (3) Peripheral auditory functions are mainly affected by aging. On the other hand, central auditory processing functions of the aged examined with word familiarity and meanings of words shows same performance as the young. These results were expected to lead the discussion for speech communication in aged society and the standardization for sound environment.

5aSC2. Effect of training using lexically easy and hard words on speech understanding in noise. Matthew H. Burk and Larry E. Humes (Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405, maburk@indiana.edu)

Older hearing-impaired adults often have difficulty understanding speech in noise, even with proper amplification. One reason for this difficulty may be a lack of practice or inability to make use of new auditory information, which has been absent due to a progressive, peripheral hearing loss over many years. An often overlooked aspect of the rehabilitation process, which could help to improve this deficit, is listener training. The goal of this study was to create a word-based training protocol which could improve speech understanding in noise when listeners are presented with new, novel stimuli outside the clinic. Previous work with word-based training using one talker showed a large training effect that generalized to novel talkers; however, sufficient generalization to novel words and sentences was lacking. The current study attempts to increase generalization by training listeners with multiple talkers and lexically hard words. Generalization to novel words, both lexically easy and lexically hard, novel talkers, and sentences, with the latter also constructed from lexically easy

and hard words, will be described. [Work supported, in part, by a research grant from NIA, R01-AG08293, awarded to the second author, and an NIH training grant.]

5aSC3. Comparison of speech intelligibility measures. Jacqueline S. Laures and Gary G. Weismer (Georgia State Univ., Atlanta, GA 30302) and (Univ. of Wisconsin-Madison, Madison, WI)

The speech intelligibility of dysarthric speakers is perceptually measured by one of the following four techniques: direct magnitude estimation with a modulus, free modulus magnitude estimation, interval scaling, and transcription. Weismer and Laures (2002) suggest that magnitude estimates may provide a more complete representation of speech intelligibility than other methods of measurement because it may be more sensitive to non-segmental aspects of speech, such as prosody. However, the empirical data supporting such a statement is quite limited. The purpose of the current study is to explore the relationship of the four different measurement techniques to determine if one approach may provide a more accurate determination of the speech intelligibility of dysarthric speakers. Twelve listeners measured the speech of six dysarthric speakers and two healthy control speakers using the four different measurement techniques. Each speaker produced three sentences twice. The sentences were presented via a loudspeaker in a sound attenuated booth. Listeners rated the sentences using the four techniques. The order of techniques used was counterbalanced. A correlation analysis revealed that the four techniques were highly related. Implications of this finding are discussed.

5aSC4. Effects of speech-rate and pause duration on sentence intelligibility in younger and older normal-hearing listeners. Akihiro Tanaka, Shuichi Sakamoto, and Yô-iti Suzuki (R.I.E.C., Tohoku Univ., Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan)

Speech-rate conversion techniques aid speech comprehension by allowing more time for perceptual and cognitive processes. However, if only the speech-rate of a telecast is converted, auditory and visual information become asynchronous. One possible method to resolve the problem is to reduce the pause durations between phrases; unfortunately, this can evoke a marked negative effect. For that reason, the present study examines the effects of the speech-rate and pause duration on sentence intelligibility. We manipulated the lengths of phrases relative to the original length (0, +100, +200, +300, and +400 ms), and the pause durations between phrases in a sentence (0, 100, 200, 300, and 400 ms). Listeners were asked to write down sentences they discerned from the noise. The intelligibility score increased in younger and older listeners when the speech signal was expanded. Regarding the pause duration, intelligibility was best when the pause duration was 200 ms in younger listeners; in older listeners, the intelligibility score was highest when the pause durations were 200 ms and 400 ms. These results provide evidence that might benefit speech-rate conversion through better use of pause duration.

5aSC5. Simulation of temporal aspects of auditory aging. Ewen MacDonald (Inst. of Biomaterials and Biomed. Eng., Rm 407 Rosebrugh, Univ. of Toronto, Toronto, ON, Canada M5S 3G9, macdone@ecf.utoronto.ca), Kathy Pichora-Fuller, and Bruce Schneider (Univ. of Toronto at Mississauga (UTM), Mississauga, ON, Canada L5L 1C6)

A jittering technique to disrupt the periodicity of the signal was used to simulate the effect of the loss of temporal synchrony coding believed to characterize auditory aging. In one experiment jittering was used to distort the frequency components below 1.2 kHz and in a second experiment the components above 1.2 kHz were distorted. To control for spectral distortion introduced by jittering, comparison conditions were created using a smearing technique (Baer and Moore, 1993). In both experiments, 16 normal hearing young adult subjects were presented with SPIN sentences in three conditions (intact, jittered, and smeared) at 0 and 8 dB SNR. When the low frequencies were distorted, speech intelligibility in the jittered

conditions was significantly worse than in the intact and smeared conditions, but the smeared and intact conditions were equivalent. When the high frequencies were distorted, speech intelligibility was reduced similarly by jittering and smearing. On low-context jittered sentences, results for young adults mimicked results found previously for older listeners with good audiograms (Pichora-Fuller *et al.*, 1995). It is argued that the jittering technique could be used to simulate the loss of neural synchrony associated with age-related changes in temporal auditory processing.

5aSC6. Comparison of hearing loss compensation algorithms using speech intelligibility measures. Meena Ramani (Dept. of Elec. and Comput. Eng., Univ. of Florida, P.O. Box 116130, Bldg. 33, Ctr. Dr. Rm. NEB 444, Gainesville, FL 32611), John G. Harris (Univ. of Florida, Gainesville, FL 32611), Alice E. Holmes (Univ. of Florida, Gainesville, FL 32611), Mark Skowronski (Univ. of Florida, Gainesville, FL 32611), and Sharon E. Powell (Univ. of Florida, Gainesville, FL 32611)

Sensorineural hearing loss includes loss of high-frequency sensitivity which results in decreased speech intelligibility. The loss cannot be compensated by inverting the audiogram because of the non-linear effects of sensorineural hearing loss (frequency smearing, decreased dynamic range, decreased time-frequency resolution). Several non-linear compensation schemes exist (Half-gain, POGO, NAL-R, Fig. 6, DSL and LGOB) and this paper provides a comparison of those using the objective Perceptual Evaluation of Subjective Quality (PESQ) score and the subjective Hearing In Noise Test (HINT). The listening tests were run on 15 unaided hearing impaired listeners as well as 15 normal hearing listeners using a simulated hearing loss algorithm. These results show marked improvement in intelligibility for the compensated speech over the normal speech for both normal and hearing impaired adults.

5aSC7. A comparative study of perceived, predicted, and measured speech intelligibility. Michael E. Hermes, Melinda J. Carney, and Dominique J. Cheenne (Dept. of Audio Arts & Acoust., Columbia College Chicago, Chicago, IL 60605)

Intelligibility metrics were obtained using a variety of methods in a gymnasium that serves as a place of worship. A word list trial, a computer model, and a computer-based %Alcons test provided the data. The results were compared in order to gauge their relative accuracy. The data from the %Alcons testing were found to be unreliable, but a direct relationship was established between the mean word list test scores and the results gathered from the computer model. This relationship allowed for a translation of the scores to %Alcons.

5aSC8. A statistical model for prediction of functional hearing abilities in real-world noise environments. Sigfrid Soli (House Ear Inst., 2100 W. 3rd St., Los Angeles, CA 90057), Chantal Laroche, Christian Giguère, and Véronique Vaillancourt (Univ. of Ottawa, Ottawa, ON, Canada)

Many tasks require functional hearing abilities such as speech communication, sound localization, and sound detection, and are performed in challenging noisy environments. Individuals who must perform these tasks and whose functional hearing abilities are impaired by hearing loss may constitute safety risks to themselves and others. We have developed and validated in two languages (American English and Canadian French) statistical techniques based on Plomps (1986) speech reception threshold model of speech communication handicap. These techniques predict functional hearing ability using the statistical characteristics of the real-world noise environments where the tasks are performed together with the communication task parameters. The techniques will be used by the Department of Fisheries and Oceans Canada to screen individuals who are required to perform hearing-critical public safety tasks. This presentation will summarize the three years of field and laboratory work culminating in the implementation of the model. Emphases will be placed on the methods

for statistical characterization of noise environments, since these methods may allow generalization of the model to a wider range of real-world noise environments. [Work sponsored by Department of Fisheries and Oceans Canada.]

5aSC9. Predicting speech intelligibility in real-world noise environments from functional measures of hearing. Christian Giguère, Chantal Laroche, Véronique Vaillancourt (Univ. of Ottawa, 451 Smyth Rd., Ottawa, ON, K1H 8M5, Canada), and Sigfrid Soli (House Ear Inst., Los Angeles, CA 90057)

In everyday life situations and in many occupational settings, speech communication is often performed in noisy environments. These environments can sometimes be very challenging, particularly for individuals impaired by hearing loss. Diagnostic measures of hearing, such as the audiogram, are not adequate to make accurate predictions of speech intelligibility in real-world noise environments. Instead, a direct functional measure of hearing, the Hearing In Noise Test (HINT), has been identified and validated for use in predicting speech intelligibility in a wide range of face-to-face speech communication situations in real-world noise environments. The prediction approach takes into account the voice level of the talker in noise due to the Lombard effect, the communication distance between the talker and the listener, a statistical model of speech perception in specific noises, and the functional hearing abilities of the listener. The latter is taken as the elevation of the individual's speech reception threshold in noise above the normative value for the HINT test. This test is available in several languages, so that language-specific needs can be addressed. The detailed approach will be presented with an emphasis placed on application examples in clinical and/or occupational settings.

5aSC10. Perceptual, acoustic, and tongue shape measures during /r/ production pre- and post-treatment using visual feedback from ultrasound: case studies of two adolescents. Marcy Adler-Bock, Barbara Bernhardt, Penelope Bacsfalvi (Dept. of Audiol. and Speech Sci., Univ. of British Columbia, 5804 Fairview Ave., Vancouver, BC, V6T 1Z3), and Bryan Gick (Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z1)

This study examined the effectiveness of using visual feedback from ultrasound in remediation of persistent /r/ errors. Ultrasound provided the participants and the clinician with a dynamic sagittal or coronal image of the tongue during speech production. The participants in this study were two adolescent boys ages 12 and 14 who were not able to produce an on-target North American /r/. Both participants had received at least one year of traditional /r/ therapy without improvement. Treatment was provided over 13 one-hour sessions using visual feedback from ultrasound. Initially, /r/ was broken down and rehearsed as individual motor targets (tongue tip, body, root); these components were then practiced in combination to produce /r/ in isolation, then in syllables, words, and phrases. Post-treatment changes in /r/ production were captured through transcription, acoustic analysis, and tongue shape measurement. Both participants /r/ productions were rated as having more tokens of on-target /r/ post-treatment. Acoustic results supported these findings with the third formant lowering post-treatment. Tongue shape measures indicated that the participants tongue shapes were more similar to the modeled /r/ tongue shape post-treatment. These case studies suggest that visual feedback as provided by ultrasound may be a useful adjunct to speech (re)habilitation.

5aSC11. The influence of severity of speech involvement on acoustic measures in dysarthria. Yunjung Kim, Gary Weismer, and Ray D. Kent (Dept. of Communicative Disord. and Waisman Ctr., Univ. of Wisconsin-Madison, Madison, WI 53705)

Several different acoustic measures have described the articulatory deficit and predicted the overall speech intelligibility deficit in speakers with dysarthria. The articulatory basis of acoustic variables that predict

speech intelligibility variations across speakers have been thought to be the ones that should be manipulated clinically for maximum therapeutic effect. For example, *F2* extent and slope are known to have strong correlations with speech intelligibility measures, across dysarthric speakers. This may suggest that within-speaker manipulation of magnitude and/or speed of articulatory movements will, if successful, result in improved speech intelligibility. Some of our previous work [Weismer *et al.*, 2001, Folia Phoniat. Logopaed.] indicates, however, that these measures may not be predictive of within-speaker fluctuations in intelligibility, but rather are indices only of across-speaker variation in severity. The large data base of dysarthric speakers at UW-Madison permits us to begin to determine how much variability exists for measures like *F2* slope in a relatively large group of speakers with homogeneous speech severity. This paper will report results of *F2* extent and slope measures for speakers with relatively mild speech involvement. If the measures are primarily tied to severity, they should not vary much within a homogeneous group of speakers. [Work supported by DC00319.]

5aSC12. Variability of jaw movement in contrastive stress production of children with and without speech delay of unknown origin. Jennell C. Vick, Lakshmi Venkatesh, and Christopher A. Moore (Dept. of Speech and Hearing Sci., Univ. of Washington, Seattle, WA 98101)

This study was designed to evaluate the control of jaw movement in contrastive stress productions of children with and without speech delay. The spatiotemporal index (STI) was used to calculate variability in jaw movement trajectories in 12 children producing three different metrical forms of CVCV syllables (trochaic, iambic, and even stress; papa, mama, and baba). The children (mean age: 3;2 years) were categorized, four in each of three groups, as having Normal Speech Acquisition (NSA), Speech Delay (SD), or both (NSA/SD) using the Speech Disorder Classification System (Shriberg *et al.*, 1997). Results replicated findings in typically developing children of a similar age reported by Goffman and Malin (1999) where iambic forms were produced with greater stability than trochaic forms as measured by the STI. Analysis using a repeated measures ANOVA revealed significant effects for contrastive stress type, speech disorder classification, and syllable identity. Results are reported in the context of vowel acoustic measures of contrastive stress. [Work supported by NIDCD R01 DC 0000822-10.]

5aSC13. Articulatory movements during vowels produced by speakers with dysarthria and normal controls. Yana Yunusova, John Westbury, and Gary Weismer (Dept. of Commun. Disord. and Waisman Ctr., Univ. of Wisconsin-Madison, 1500 Highland Ave., Madison, WI 53705)

The central goal of the study was to provide a quantitative description of contributions of the jaw, lower lip, tongue blade and tongue dorsum to vowel productions, and to determine patterns of interarticulatory interactions between movements. Kinematic and acoustic signals were collected using the x-ray microbeam. Thirty-four speakers, 12 with dysarthria due to Parkinson disease, 7 with amyotrophic lateral sclerosis, and 15 normal controls, were recorded reading sentences at a comfortable speech rate. Ten CVC words, each containing one of the English vowels /i, I, u, a, ae, o/ carrying primary stress, were selected for analysis. Each fleshy point trajectory was characterized by marker positions at vowel onset and offset, and the moment when speed was lowest. Measures of distance traveled, time to and from the moment of minimum speed, and peak and average movement speed were employed. Movement characteristics, and associations between movements, were compared for different vowels, contexts, speakers and groups. Results are reported for vowels and vowel groups (e.g., lax versus tense), averaged separately by contexts for speaker groups. The data speak to previous claims that speakers with dysarthria exhibit evidence of discoordination in speech movements relative to normal performance. [Work supported by NIDCD Award R01 DC003723.]

5aSC14. Spectral contributions to intelligibility of sentences with flattened fundamental frequency. Peter J. Watson and Robert S. Schlauch (Dept. of Speech-Lang-Hearing Sci., Univ. of Minnesota, 164 Pillsbury Dr. S.E., Shevlin 115, Minneapolis, MN 55455, pjwatson@umn.edu)

Recently, the contribution of fundamental frequency (F_0) variation to speech intelligibility in background noise has been studied [J. Laures and G. Weismer, *JSLHR* **42**, 1148–1156 (1999)]. The process used for this analysis is to flatten the frequency contour at the average F_0 . Results show that sentences with flattened F_0 are less intelligible than those with natural F_0 variation. However, this technique may reduce the prominence of formant peaks because it excludes F_0 s that are below the average flattened frequency. As noted by Laures and Weismer (1999), eliminating lower F_0 s from the analysis results in wider spacing between the harmonics and the available energy to excite the vocal tract resonances is diminished. This study examined the effect of flattening at the average and lowest F_0 of each sentence to intelligibility. Twenty subjects listened to sentences in a continuous noise background. Sentences were equated for root-mean square energy. Results showed that the low-flattened condition was significantly more intelligible than the average-flattened condition. However, a condition with natural F_0 variation was more intelligible than the two flattened conditions. These findings show that the frequency of flattening does contribute to intelligibility but that natural F_0 variation appears to contribute to a greater degree.

5aSC15. Comparison of cues in auditory selective and divided attention. Jae hee Lee and Larry E. Humes (Indiana Univ. at Bloomington, Bloomington, IN 47405)

This study investigated auditory selective and divided attention in ten normal-hearing listeners. Subjects listened to two simultaneous sentence-like utterances from the Coordinate Response Measure (CRM) speech corpus at 90 dB SPL in dichotic or monotic listening conditions. This corpus has the following format, Ready (call sign), go to (color) (number), now, with all possible combinations of call signs (8), colors (4) and numbers (8). In all conditions, the listener identified the color-number coordinates spoken by the target talker. The target talker was identified by a cue provided either prior to (selective attention) or following (divided attention) the presentation of the two simultaneous utterances. In the first experiment, either the ear or the call sign served as the cue for the target talker in monotic or dichotic listening conditions. In the second experiment, talker gender served as the cue in monotic listening conditions. Across both experiments, performance for selective attention was superior to that for divided attention. Performance for ear or gender cueing was better than for call sign. Analysis showed that 80% of errors were due to intrusions from the competing utterance, supporting an informational (attention), rather than energetic, form of masking. [Work supported, in part, by NIH R01 AG08293.]

5aSC16. Effects of high intensity on recognition of low- and high-frequency speech in noise. Van Summers and Mary Cord (Army Audiol. & Speech Ctr., Walter Reed Army Medical Ctr., Washington, DC 20307-5001)

For listeners with normal hearing (NH), speech recognition scores often decrease when intensities exceed moderate levels (rollover is observed). It is currently unclear whether the factors leading to rollover in NH listeners also limit performance for hearing-impaired (HI) listeners at high sound levels. This study aimed at clarifying the stimulus conditions most clearly associated with rollover and whether rollover effects are similar for NH listeners and listeners with mild hearing impairment. In Stage 1, NH and HI listeners heard digitally-filtered sentences and adaptive procedures were used to determine high- and low-frequency bandwidths yielding 70%-correct word recognition in quiet at moderate levels. In Stage 2, broadband and band-limited stimuli (based on the high-, and low-frequency passbands measured in quiet in Stage 1) were tested at

moderate levels in background noise. Noise levels were varied adaptively to determine signal-to-noise levels supporting 30%-correct recognition. Stimulus conditions leading to criterion performance at moderate presentation levels in Stage 2 were then retested at higher levels in Stage 3. NH listeners showed larger and more consistent rollover effects for high-frequency than for low-frequency or broadband stimuli. The results for HI listeners showed greater variability but also indicated clear rollover effects for high-frequency stimuli at high levels.

5aSC17. The effect of auditory feedback alterations on the speech quality of hearing aid and cochlear implant users. Dragana Barac-Cikoja and Leslie Klein (Gallaudet Univ., 800 Florida Ave. NE, Washington, DC 20002)

The speech of seven hearing aid (HA) users with severe-profound hearing loss and six cochlear implant (CI) users was recorded as they read the Rainbow Passage while their speech feedback was either spectrally altered in real-time, completely masked by multi-talker babble noise, or unaltered. Spectral alterations were implemented by filtering the speech signal into either one or four frequency bands, extracting their respective amplitude envelope(s), and amplitude-modulating the corresponding noise band(s). While the single-band condition provided only coarse information about the speech rhythmic structure, the four-band noise signal remained intelligible. Auditory feedback was presented via insert earphones to the HA users, and via the auxiliary jack (with the headpiece microphone silenced) to the CI users, at the participants' most comfortable listening level. The quality of the recorded speech (separated into individual sentences) was assessed using a 2IFC procedure. For each combination of the experimental conditions, six judges selected the more natural-sounding utterance in a pair. Preference scores were calculated for each of the four feedback conditions and statistically tested. HA and CI group differed in how feedback intelligibility affected the speech quality. Possible acoustic correlates of the perceived differences will be discussed.

5aSC18. Articulatory and acoustic measurements of vowels in hearing impaired speakers following treatment. Penelope Bacsfalvi and Barbara Bernhardt (School of Audiol. and Speech Sci. (SASS), UBC, Vancouver, BC, Canada V6T 1Z3, Penelope@audiospeech.ubc.ca)

The purpose of this study was to examine the relationships between ultrasound tongue shapes, electropalatography (EPG) contacts, formant data, and the perceptual data in the vowels of three hard of hearing adolescents. This pilot study examines the pre- and post-therapy speech of the participants in a 6 week group therapy programme with ultrasound and EPG technologies. Before treatment, participants showed particular difficulty with high vowels and the tense-lax distinction. Recordings were made of 2 males and 1 female with ultrasound and EPG. Three types of measurements were made; formant measurements, EPG tongue palate contacts, and perceptual judgements by experienced speech-language pathology researchers. Initial analysis values showed a change in the direction of EPG contact patterns, and perceptual categories. F_1 values decreased and F_2 values increased, moving in the direction of typical formant frequencies found in hearing people. Preliminary results of this study support a lingual-articulatory approach to treatment.

5aSC19. The role of fundamental frequency (F_0) contours in the perception of speech against interfering speech. Christine Binns and John F Culling (School of Psych., Cardiff Univ., Tower Bldg., Park Pl. Cardiff, CF10 3AT, UK)

A set of experiments investigated the effect of the F_0 contour on speech intelligibility against interfering sounds. Speech Reception Thresholds (SRTs) were measured in dB for sentences with different manipulations of their F_0 contours. These manipulations involved either a scaled reduction in F_0 variation, or the complete inversion of the F_0 contour. Against speech-shaped noise, a flattened F_0 contour did not have a sig-

nificant impact on the intelligibility of speech compared to a normal *F0* contour: the SRT for the flattened *F0* contour being only 0.7 dB higher. The SRT for the inverted contour, however, was 1.6 dB higher than for the normal *F0* contour. When the sentences were played against a single-talker interferer, the overall effect was greater, with a 2.1 dB difference between the normal and flattened conditions, and 3.3 dB between the normal and inverted *F0* contours. There was found to be no effect of altering the *F0* contour of the interferer, indicating that any abnormality of the *F0* contour serves to mask the intelligibility of the target speech, but does not alter the effect of the interfering speech. Future research will investigate the impact on speech intelligibility of similar manipulations of a low-pass-filtered *F0* contour.

5aSC20. Perceptual strategies for identifying vowels produced by speakers with dysarthria. Kate Bunton (Dept. of Speech, Lang., and Hearing Sci., Univ. of Arizona, P.O. Box 210071, Tucson, AZ 85721)

It has been previously reported that changes to the fundamental frequency contour of vowels influences their identification [e.g., H. Traunmuller, *J. Acoust. Soc. Am.* **69**, 1465–1475 (1981)]. Similar changes in vowel identification have been demonstrated for some listeners when the *F0* contour of a dysarthric speaker has been modified (e.g., flattened or enhanced) [Bunton (2004)]. These listeners appear to rely on *F0*–*F1* space for identification of vowel height; this strategy is consistent with a perceptual model of vowel recognition described by A. Syrdal and H. Gopal [*J. Acoust. Soc. Am.* **79**, 1086–1100 (1986)]. Other listeners, however, appear to utilize different acoustic cues to identify a given vowel. This study focused on defining those perceptual strategies. Multi-dimensional plots of acoustic cues likely to influence vowel perception are used to illustrate perceptual strategies that may have been used by these listeners for identification of vowels in the speech of individuals with dysarthria associated with Parkinson disease. [Work supported by NIH R03 DC005902.]

5aSC21. Perception of coarticulation in gated words by dyslexic and non-dyslexic children. Patricia Keating (UCLA, Los Angeles, CA 90095-1543), Frank Manis, Jennifer Bruno, and Jonathan Nakamoto (USC, Los Angeles, CA 90089-1061)

In an auditory word gating task, listeners are presented with portions of words, and try to identify these acoustic fragments as lexical items. It has been shown that children need more acoustic information than adults to succeed, and that dyslexic children can require more information than other children. It is also known that adults can use early acoustic cues to identify features of upcoming segments. For example, in English anticipatory vowel nasalization implies that a nasal consonant will follow, and similarly anticipatory lateralization. Our study asked whether dyslexic children are impaired in their perception or use of such anticipatory coarticulatory information. Successive gates from test words that ended in a nasal (8 words) or /l/ (4 words), or control words with final oral stops, were blocked by length and presented to 26 dyslexic and 26 non-dyslexic children. Responses were audiorecorded and later transcribed; responses were scored re both the full word and the nasality/laterality of the final consonant. Dyslexics as a group required more gates to identify nasality, but not whole words. Language-impaired dyslexics required more gates to identify whole words as well as nasality. Performance on the gating task significantly predicted reading ability when age and IQ were controlled.

5aSC22. Phonetic labeling along a formant transition duration continuum in children with specific language impairment. Harvey M. Sussman, Elizabeth Burlingame, Ronald B. Gillam (Dept. of Commun. Sci. & Disord., Univ. of Texas, Austin, TX 78712), and Jessica F. Hay (Univ. of Texas, Austin, TX 78712)

Fifteen children diagnosed with specific language impairment (SLI) and fifteen typically developing (TD) children were tested for identification performance on two synthetic speech continua varying in formant transition durations (FTDs). One continuum varied from [ba]-to-[wa] and

the other varied from [da]-to-[ja]. Several dependent measures were derived based on *d'* from signal detection theory. These measures were used to assess category boundaries and to indirectly assess sensitivity to phonetic changes in labeling category tokens along each continuum. The SLI group showed less consistent identification performance along the [ba]-[wa] series relative to the TD group, as well as reduced sensitivity to phonetic changes along the continuum. On the [da]-[ja] continuum, the SLI group revealed less consistent identification performance on the short FTD end, but similar identification levels to the TD group at the long FTD end. The overall results support the view that children with SLI reveal a deficiency in the processing of speech sounds at the level of segmental identity.

5aSC23. Classification of dysarthric and non-impaired speech based on prosodic features. Greg Kochanski (Oxford Univ., 41 Wellington Square, Oxford OX1 2JF, UK, gpk@kochanski.org) and Rupal Patel (Northeastern Univ., Boston, MA 02115)

Prosodic differences between dysarthric and healthy speakers were studied. Six acoustic properties that are plausibly more influenced by suprasegmental aspects of speech (e.g., emphasis) than the segmental details of the words were measured. The time course of these properties were analyzed over each utterance by fitting Legendre Polynomials. The resultant Legendre coefficients were then fed to linear- and quadratic-discriminant classifiers. All of the six properties were individually capable of distinguishing dysarthric speech from healthy speech. Based on one acoustic property measured over a single short sentence, we could correctly classify a speaker as healthy or dysarthric 55–75% of the time, depending on the acoustic property used. More complex classifiers that used all the acoustic properties correctly classified the speaker 97% of the time based on nine utterances. The strongest difference between normal and dysarthric speech was in loudness. Dysarthric speakers did not reliably produce the loudness patterns associated with stressed syllables. They also had a wider range in amplitude, greater voicing variability, smaller excursions of fundamental frequency, and less final lengthening compared to healthy speakers. The classification we demonstrated may be extended to become a graduated measurement of severity, thereby contributing to diagnostics and intervention in dysarthria.

5aSC24. Effects of hearing impairment on the perception and neural representation of time-varying spectral cues. Ashley W. Harkrider, Patrick N. Plyler, and Mark S. Hedrick (Dept. of Audiol. and Speech Pathol., Univ. of Tennessee, 457 South Stadium Hall, Knoxville, TN 37996)

Differences in phonetic boundaries versus normal controls suggest that listeners with hearing impairment (HI) have difficulty categorizing stop consonant place of articulation based solely on the dynamic spectral information present in the second formant transition (*F2*), even when the stimuli are amplified. This may be due to a degraded ability of the central auditory nervous system to process time-varying spectral cues despite ensuring overall audibility. However, increasing the overall level of the stimuli may not result in improved audibility of *F2*. To determine if spectral shaping of *F2* improves performance of listeners with HI, psychometric functions and *N1*-*P2* cortical responses were compared in 10 older listeners with normal hearing versus 10 older listeners with HI. Stimuli were synthetic consonant-vowels along a /ba/-/da/-/ga/ place-of-articulation continuum in an unshaped and shaped condition. Generally, behavioral and *N1*-*P2* results indicate that, with shaping, categorization of /d/ and /g/ improves. These findings suggest that enhanced audibility of *F2* through spectral shaping does improve perception of stop consonant stimuli. However, categorical boundaries for the individuals with HI are shifted lower in frequency with shaping for all phonemes versus normal controls, indicating that enhancing audibility improves but does not completely normalize categorization performance.

5aSC25. Effect of two-band dichotic listening for hearing impaired listeners. Shuichi Sakamoto, Atsunobu Murase, Yōiti Suzuki (Res. Inst. of Elect. Comm./ Grad. School of Information Sci., Tohoku Univ., 2-1-1 Katahira, Aoba-ku, Sendai, Miyagi, Japan, saka@ais.riec.tohoku.ac.jp), Tetsuaki Kawase, and Toshimitsu Kobayashi (Tohoku Univ., Aoba-ku, Sendai, Miyagi, Japan)

The increase of the upward spread of masking is a phenomenon that is typically observed among sensorineural hearing-impaired listeners. To resolve this problem, dichotic listening, by which an input speech spectrum is split into two complementary parts and is presented dichotically, seems effective to reduce masking between contiguous frequency bands. This study examines effects of simple two-band dichotic listening with a cut-off frequency around and between the typical first and second formant frequencies of the preceding vowel. We measured speech intelligibilities in both quiet and noisy environments (S/N 4 and 0 dB). Three types of vowel-consonant-vowel nonsense monosyllables, of which preceding vowels were /a/, /i/, and /u/, were used as speech stimuli. Results showed that this dichotic processing was effective, especially in relatively high S/N conditions. Moreover, the best dividing frequency was dependent on the preceding vowel. When /a/-consonant-vowel was used, the best dividing frequency was 1.0 kHz (around F_1 of Japanese vowel /a/), whereas the best dividing frequency was 0.8 kHz (between F_1 and F_2 of Japanese vowel /u/) when the /u/-consonant-vowel was used.

5aSC26. Signal to noise ratio loss and consonant confusions. Yangsoo Yoon and Jont B. Allen (Univ. of Illinois, Speech and Hearing, 901 s sixth, Champaign, IL 61820, yyoons5@uiuc.edu)

Previous SNR loss (also called speech loss) studies showed that (1) SNR loss cannot be predicted from audiometric measures, (2) 40% of hearing aids wearers have 5 dB SNR loss or greater, and (3) SNR loss influences speech intelligibility significantly. These showed SNR loss to be important in speech recognition, but they do little, or no to illuminate the nature of consonant confusion, resulting from SNR loss. Thus, the goal of the current study was to investigate the effect of SNR loss on 16 consonants recognition in hearing impairment as a function of SNR. Confusion matrix data were collected and analyzed, and Fletcher's AI was calculated from the SNR. These two measures were utilized (1) to determine how SNR loss was related to the event loss, (2) to test whether clustering of syllables in terms of consonant confusions was complied with SNR loss, and (3) to compare PI functions obtained from subjects and AI model. The results show that the degree of consonant confusion varies, but members of consonants confused with target sound above chance level are similar, as a function of SNR loss and SNR. It suggests that SNR loss limits recognition for specific consonants, even in noise.

5aSC27. Driving performance and auditory distractions. Elzbieta B. Slawinski, Jane F. MacNeil, Mona Motamedi, Benjamin R. Zendel (Psych. Dept., Univ. of Calgary, 2500 Univ. Dr., Calgary, AB, Canada T2N 1N4), Kirsten Dugdale, and Michelle Johnson (Univ. of Calgary, Calgary, AB, Canada T2N 1N4)

Driving performance depends on the ability to divide attention during different tasks. In spite of the fact that driving abilities are associated with visual stimulation, driving performance depends on attention to stimulation and/or auditory distraction. Research shows that listening to the radio is a principal auditory distracter during the time of driving (Brody, 2002). In the laboratory a few experiments were conducted on the auditory distraction (e.g., music, stories) and signal processing by young and older drivers. Results show that older subjects involved in listening to the stream of information (independent of the hearing status) require higher intensity of the auditory stimulation than younger drivers. It was shown that cognition plays a role while listening to auditory stimuli. Moreover, it was demonstrated that driving performance was influenced by the type of per-

formed music. A portion of these experiments and their results were presented at the Annual Meetings of CAA in 2002 and 2003 as well as being published in the Journal of Psychomusicology **18**, 203–209. Complete results of the experiments will be discussed.

5aSC28. Speech intelligibility index calculations in light aircraft cabin during flight. Tino Bucak and Ernest Bazijanac (Dept. of Aeronautics, Faculty of Transport and Traffic Eng., Univ. of Zagreb, Croatia)

High levels of cabin noise in small general aviation aircraft significantly deteriorate the quality of speech communications and potentially endanger the safety of flight. Several ground and inflight cabin noise measurements on new generation Cessna 172R were made during various phases of flight. The results are analyzed and used for Speech Intelligibility Index (SII) calculations, in order to quantify the influence of cabin noise on speech communications between crew members.

5aSC29. A detailed study on the effects of noise on speech reception. Tammo Houtgast and Finn Dubbelboer (VU Univ. Medical Ctr., Amsterdam, The Netherlands)

The effect of adding continuous noise to a speech signal was studied by comparing, for a series of quarter octave bands, the band output for the original speech and for the speech-plus-noise. Three separate effects were identified. (a) Average envelope-modulation reduction: the original intensity-envelope is, on average, raised by the mean noise intensity, resulting in a reduction of the original modulation index. (b) Random instantaneous envelope fluctuations: on an instantaneous basis, the speech-plus-noise envelope shows random variations, caused by the stochastic nature of the noise, and by the instantaneous changes in the phase relation between the speech and the noise. (c) Perturbations of the carrier phase: in the band output carrier signal the addition of the noise causes random phase changes. By applying signal processing techniques, we were able to either include or exclude each of these three effects separately. The results of intelligibility measurements indicated the following order of importance of the three different effects: (1) the average envelope-modulation reduction, (2) the perturbation of the carrier phase, and (3) the random envelope fluctuations. The results will be discussed in the light of modeling and enhancing (noise suppression schemes) speech reception in noise.

5aSC30. Speech rate characteristics in dysarthria. Kris Tjaden, Geoff Greenman, Taslim Juma, and Roselinda Pruitt (Dept. of Communicative Disord., Univ. at Buffalo, 122 Cary Hall, 3435 Main St., Buffalo, NY 14214, tjaden@acsu.buffalo.edu)

Speech rate disturbances are pervasive in dysarthria, with some reports suggesting that up to 80% of speakers with dysarthria exhibit speech rates that differ from neurologically normal talkers. The contribution of articulation time and pause time to the overall impairment in speech rate is not well understood. Studies investigating speech rate characteristics in dysarthria also tend to focus on reading materials, yet there is reason to suspect that the higher cognitive load of conversational speech may impact speech rate characteristics differently for individuals with impaired speech motor control, and neurologically normal talkers. The current study will report speech rate characteristics for both a reading passage and conversational speech produced by individuals with dysarthria secondary to Multiple Sclerosis (MS), individuals with dysarthria secondary to Parkinson Disease (PD) and healthy controls. The manner in which speech rate, articulation rate, and pause characteristics differ for speakers with dysarthria and healthy controls will be examined. The contribution of articulation time and pause time to overall speech rate also will be studied and compared for the reading passage and conversational speech. [Work supported by NIDCD R01DC04689.]

Session 5aSP**Signal Processing in Acoustics: Smart Acoustic Sensing for Land-Based Surveillance**

Brian Ferguson, Chair

*Defense Science and Technology Organization Maritime Systems Div., Pyrmont 2009, Australia***Chair's Introduction—8:30*****Invited Papers*****8:35****5aSP1. Acoustic methods for tactical surveillance.** Brian G. Ferguson and Kam W. Lo (Defence Sci. and Technol. Organisation, P.O. Box 44, Pyrmont, NSW 2009, Australia)

Smart acoustic sensor systems can be deployed for the automatic detection, localization, classification and tracking of military activities, which are inherently noisy. Acoustic sensors are appealing because they are passive, affordable, robust, and compact. Also, the propagation of sound energy is not limited by obstacles that block or obscure the clear line of sight that is required for the effective operation of electromagnetic systems. Methods, with examples, for extracting tactical information from acoustic signals emitted by moving sources (air and ground vehicles) are provided for both single sensor and multiple sensor configurations. The methods are based on processing either the narrowband or broadband spectral components of the sources' acoustic signature. Weapon firings generate acoustic impulses and supersonic projectiles generate shock waves enabling source localization and classification by processing the signals received by spatially-distributed sensors. The methods developed for land-based acoustic surveillance using microphone data are also applied to hydrophone data for passive acoustic surveillance of the underwater environment.

9:05**5aSP2. Autonomous acoustic sensing on mobile ground and aerial platforms.** Tien Pham and Nassy Srour (US Army Res. Lab., 2800 Powder Mill Rd., Adelphi, MD 20783-1197)

Acoustic sensor systems on the ground and/or in the air can be used effectively for autonomous and remote intelligence surveillance and reconnaissance (ISR) applications. Acoustic sensors can be used as primary sensors and/or secondary sensors to cue other higher-resolution sensors for detection, tracking and classification of continuous and transient battlefield acoustic events such as ground vehicles, airborne aircraft, personnel, indirect fire, and direct fire. Current collaborative research activities at ARL in acoustic sensing from mobile ground platforms such as HMMWVs and small robotic vehicles [P. Martin and S. Young, Proc. of SPIE Defense & Security Symposium, 2004] and from aerial platforms such as UAVs and balloons [Reiff, Pham *et al.*, Proc. of the 24th Army Science Conference, 2004] demonstrate practical performance enhancements over fixed ground-based platforms for a number of ISR applications. For both mobile ground and aerial platforms, self-generated noise (flow noise and platform noise) is problematic but they can be suppressed with specialized windscreens, sensor placement, and noise cancellation technology. Typical acoustic detection and processing results for mobile platforms are compared and contrasted against fixed ground-based platforms.

9:35**5aSP3. Ferret and its applications.** Jacques Bedard (Defence R&D Canada—Valcartier, 2459 Pie XI North, Val-Belair, QC, Canada G3J 1X5)

Ferret is an acoustic system that detects, recognizes and localizes the source and direction of small arms fire. The system comprises a small array of microphones and pressure sensors connected to a standard PC-104 computer that analyzes, displays, reports and logs the parameters of a recognized shot. The system operates by detecting and recognizing the ballistic shock wave created by the supersonic bullet, combined with the muzzle blast wave propagating from the weapon. The Canadian Land Force Test and Evaluation Unit evaluated a vehicle-mounted version of the system and recommended deployment of the system during peacekeeping missions. The system is the result of a collaborative effort between Defence R&D Canada and MacDonald Dettwiler and Associates. This presentation describes the hardware and software components of the system along with the current and future applications of the system.

Contributed Papers

10:20

5aSP4. Acoustic self-localization of a wireless sensor network. Peter L. Schmidt, Stephen M. Williams, and Kenneth D. Frampton (Vanderbilt Univ., 2400 Highland Ave., Nashville, TN 37212)

One of the biggest challenges to the field of wireless sensor networks is self-localization: that is the determination of the relative and absolute coordinates of each sensor node in the network. Previous work has been done to locate hydrophone arrays. However sensor networks have some unique constraints that make this more challenging. A typical application would involve the distribution of hundreds or thousands of sensor nodes over an area either by hand, airdrop or other means. One of the primary constraints on such a system is that centralized processing of self-localization data may be prohibitively complex. Furthermore, the data may be incomplete, contain reflected path events, and may be subject to other mission specific constraints. Therefore, a distributed computational scheme has been developed to solve acoustic time-of-arrival equations. *A priori* information about some sensor locations and user triggered source localization events are used along with a regularized inversion solution. Results of this will be presented based on both small scale experiments and larger systems simulations. Limits of the types of *a priori* information required for accurate results are detailed, along with studies of the accuracies obtained using various distributed calculations parameters.

10:35

5aSP5. Distributed source localization in a wireless sensor network. Stephen M. Williams, Peter L. Schmidt, and Kenneth D. Frampton (Vanderbilt Univ., 2400 Highland Ave., Nashville, TN 37212)

This work concerns experimental implementation of distributed acoustic source localization with a wireless sensor network. The envisioned application is the distribution of hundreds or thousands of small, semi-disposable, wireless sensor nodes (possibly by airdrop). These nodes would form an ad-hoc communications network and monitor and track acoustic events within the array. Centralized processing of such data would be prohibited by the excessive communications and complexity of calculations. Furthermore, localization with a network of randomly placed sensors is not suited to use of traditional array theories due to irregular placement of sensors and the unknown sensitivity and phase relationships between sensors. Therefore, a fully distributed localization algorithm was developed in which nodes organize themselves into groups which collaborate to locate sources. Acoustic event time-of-arrival information is shared among group members and the source locations were determined using Tikhonov regularized inversion. Small scale localization experiments (with 10 sensor nodes) were conducted to validate the performance of the method. Results are discussed, along with limitations of the method discovered during conduct of the experiments. In addition, a comparison is made between this method and traditional matrix solution methods.

10:50

5aSP6. Particle filters for tracking a vehicle through a field of acoustic directional frequency and recording (DIFAR) sensors. Anton J. Haug (The MITRE Corp., 7515 Colshire Dr., McLean, VA 22102, ahaug@mitre.org)

For nonlinear/non-Gaussian processes, Bayesian approaches to tracking require integration over probability density functions, which cannot be accomplished in a closed form. Recently, numerical Monte Carlo “particle filter” integration techniques have been developed and applied to tracking problems. However, multidimensional application examples of these tech-

niques have been limited because they require an analytic expression for the likelihood function, which is usually not easy to obtain when the noise is non-Gaussian and not additive. Tracking a vehicle through a field of acoustic DIFAR sensors is such an application. Although the vehicle dynamic model can be linear with additive Gaussian noise, the observation model is inherently nonlinear with observations that contain embedded non-Gaussian noise. For the DIFAR vehicle tracking problem, we will consider large time-bandwidth signals, develop both the dynamic and observation models, and show how the Gaussian additive noise on each DIFAR channel results in non-Gaussian noise embedded in the bearing observations. The likelihood function will then be presented and the performance of several particle filter trackers will be compared with that of an Extended Kalman Filter.

11:05

5aSP7. Source localization in complex scattering environments using random amplitude and bearing information. D. Keith Wilson, Mark L. Moran, and Roy Greenfield (USACE ERDC-CRREL, 72 Lyme Rd., Hanover, NH 03755-1290, D.Keith.Wilson@erdc.usace.army.mil)

By scattering sound waves, atmospheric turbulence and other outdoor objects (such as building and trees) induce random fading and angle-of-arrival variations in signals received by small-baseline microphone arrays. This study addresses the relative utility of signal amplitude and bearing estimates when such variations are present. Source-localization simulations are performed for four idealized statistical scattering models: no scattering, weak turbulent scattering, strong turbulent scattering, and diffuse (multiple) scattering. Each of these cases is considered with low, moderate, and high SNR. It is found that bearing information provides highly accurate source localization when SNR is high and scattering is negligible. However, when the SNR is low and/or there is significant scattering, the bearing information loses its utility while the amplitude information remains relatively robust. Algorithms are considered that attempt to assess the relative reliability of the bearing and amplitude information, and subsequently weight signal features to provide the most satisfactory results. The simulations also confirm previous analyses suggesting that, for unfavorable propagation conditions and SNR, Cramer-Rao lower bounds predict substantially better performance than is obtainable in practice.

11:20

5aSP8. A method for locating nearby sources. John V. Olson, Kenneth Arnault, and Curt. A. S. Szuberla (Geophysical Inst., Univ. of Alaska, Fairbanks, AK 99775)

A fast method for locating nearby sources based solely on the time-of-flight information is described. The time delays for a signal passing between all sensor pairs is determined using cross-correlations and a vector of time delays is constructed. In the noise free case, where the speed of sound is constant, each point in the plane is associated with a unique value of the time delay vector. The method we present uses a fast, simplex-based search of the plane for a source location by minimizing the difference between the vector associated with the candidate location on the plane and the value estimated from cross-correlations. The search takes place over the three dimensional space that includes two coordinates in the plane and the propagation speed. The starting point for the search is constructed from an analytic fit of circular wave fronts to groups of sensors within the array. This method has been useful in identifying near-field sources in the 153US and 155US CTBT/IMS infrasound arrays. Examples of the locations we have determined along with a description of the statistical confidence limits associated with the method will be presented.

Session 5aUW

Underwater Acoustics: Scattering and Reverberation

Purnima Ratilal, Chair

Northeastern Univ., Electrical and Computer Engineering, 409 Dana Research Ctr., Boston, MA 02115-5000

Contributed Papers

8:30

5aUW1. Identification of strong scatterer locations on the Malta plateau. Mark K. Prior (NATO Undersea Res. Ctr., Viale San Bartolomeo 400, 19138, La Spezia, Italy, prior@saclantc.nato.int)

Reverberation data gathered in the Malta Plateau region of the Mediterranean Sea using a Low-Frequency Active Sonar (LFAS) and a cardioid receiver array are presented. A processing strategy involving normalization, geo-referencing and averaging over pings is described and is shown to highlight the locations of strong scatterers and to provide a first-order estimate of their strength. A second process of calculating the ratio of local standard deviation and mean values is shown to reveal scatterer locations and to provide an indication of whether the probability density function of the reverberation is well-described by the Rayleigh distribution. The characteristics of the observed scatterers are revealed by comparison with independent sidescan, echo sounder and boomer data. Areas of high reverberation are identified as being associated with wrecks, oil production platforms, rock outcrops and variations in local sediment properties.

8:45

5aUW2. Range-dependent 3D scattering and reverberation in the continental shelf environment from biology, geology and oceanography. Purnima Ratilal (Northeastern Univ., 409 Dana Res. Ctr., Boston, MA 02115), Sunwoong Lee, Yisan Lai, Tianrun Chen, Deanelle Symonds (MIT, Cambridge, MA 02139), Ninos Donabed (Northeastern Univ., Boston, MA 02115), and Nicholas Makris (MIT, Cambridge, MA 02139)

Several unified scattering and reverberation models were developed in support of the ONR Acoustic Clutter Program. They include a range-dependent model based on the parabolic equation that can be used to efficiently model scattering from a random spatial distribution of random targets that obey the sonar equation in the waveguide and a similar but range-independent waveguide model based on normal modes for scattering from extended objects. Both these models are bistatic and fully 3D, and the latter model also accounts for modal coupling between propagation and scattering caused by extended objects in the waveguide. These models are applied to examine both coherent and diffuse scattering measured after beamforming and match-filtering on an array from schools of fish, plankton, volume inhomogeneities in the sea bottom, roughness on the seafloor, extended seafloor and sub-bottom features such as river channels and reflective strata, and internal and surface waves. We provide a review of the dominant sources of clutter as well as background reverberation for long range active sonar based on comparison of model predictions with measured data from the Acoustic Clutter Experiments of 2001 and 2003.

9:00

5aUW3. Dominant source of background reverberation during the Acoustic Clutter 2003 experiment. Ninos Donabed, Purnima Ratilal (Northeastern Univ., 409 Dana Res. Ctr., Boston, MA 02115), and Nicholas Makris (MIT, Cambridge, MA 02139)

Reverberation data from the Acoustic Clutter 2003 experiment on the New Jersey Continental shelf are analyzed to provide estimates of the background scattering strength over wide areas of the environment as a

function of time. We corrected the acoustic data in the 390–440 Hz range for two-way transmission loss modeled using the range-dependent parabolic equation, the spatially varying beam pattern of the receiving array, and source level to obtain estimates of the scattering strength per unit area of the environment. Data from regions of space devoid of strong clutter were statistically analyzed to provide a temporal and spatial characterization of the background scattering strength. They were then compared to scattering strength levels deduced from modeling from a variety of scatterers including schools of fish, the sea bottom, internal waves and surface waves. Analysis indicates that the dominant sources of background reverberation can be either low density fish schools or the sea bottom.

9:15

5aUW4. Higher moment estimation for shallow water reverberation. Kevin LePage (Naval Res. Lab., Code 7144, 4555 Overlook Ave. SW, Washington, DC 20375)

Shallow water reverberation is characterized by clutter which causes false alarms in sonar systems. In previous work we have estimated the non-Rayleighness of shallow water reverberation time series predictions obtained by the R-SNAP model. Here we extend the R-SNAP model to the closed form estimation of the second moment of reverberation intensity, allowing the estimation of pdf fits for the two parameter K distribution as a function of system and channel characteristics and time after shot. This work is intended to theoretically guide the selection of clutter-robust system characteristics. [Work supported by ONR.]

9:30

5aUW5. Statistical characterization of sonar-like clutter observed on the STRATAFORM during the 2003 Acoustic Clutter Experiment in the 400–1500 Hz region. John Preston and Douglas Abraham (Appl. Res. Lab., The Penn State Univ., State College, PA 16804)

In 2003 ONR sponsored the Acoustic Clutter Experiment to study shallow water scattering and clutter in the STRATAFORM area off New Jersey. Sources were bistatic coherent pulses from a vertical array. The receiver was the Five Octave Research Array (used horizontally). The STRATAFORM is known to have benign surface morphology but contains many buried river channels and other sub-surface horizons. MIT researchers have shown fish to be a primary source of the observed clutter and reverberation. The K -distribution's shape and scale parameters have been shown to be useful in describing non-Rayleigh behavior. Statistical characterization is presented as a function of location. The "bandwidth" effect is shown where the shape parameter first decreases inversely proportional to bandwidth but then increases in a trend back toward the Rayleigh distribution at higher bandwidths. The shape parameter estimates are well fit by an elongated patch model of Abraham and Lyons. Differences between the 2003 and 2001 data taken in the same area are discussed. It is believed that the sound speed profiles in 2003 lead to more bottom interaction than in 2001 producing a larger time spread in the multipaths leading to observed differences. [Work supported by ONR Code 32, Grant N00014-05-1-0156.]

9:45

5aUW6. Comparison of measured buried target backscatter levels with high-order predictions: Rippled interface effects. Raymond Lim (Naval Surface Warfare Ctr.-Panama City, Code R21, 110 Vernon Ave., Panama City, FL 32407), Joseph L. Lopes, and Gary S. Sammelmann

Controlled sonar measurements with buried spheres and cylinders have demonstrated diffraction by bottom ripple can significantly improve target detection performance with sonar operated at grazing angles shallower than the critical grazing angle. Furthermore, buried target scattering models based on low-order perturbation theory to handle sound transmission through bottom roughness confirm that a significant enhancement in detection performance can be expected. But comparisons between measurements and model predictions have not always produced good agreement. As the bottom ripple height increases, discrepancies have been noted in comparisons with the predicted backscatter levels from a buried sphere, suggesting a failure of low-order perturbation theory to properly account for high-amplitude roughness effects. Recently, a recursive algorithm to generate arbitrary-order perturbation corrections for transmission and reflection through an idealized sinusoidal ripple was formulated and incorporated into our existing buried target scattering codes. The resulting higher-order predictions indicate previous discrepancies were not necessarily due to the neglect of high-order corrections. Calculations show model inputs such as bottom attenuation can influence backscatter levels significantly and measurement of these parameters may need to be performed more carefully. We demonstrate this through a presentation of data and model comparisons. [Work supported by ONR and SERDP.]

10:00–10:15 Break

10:15

5aUW7. Calculation of scattering from underwater targets using the equivalent source technique. Ahmad T. Abawi and Michael B. Porter (Heat, Light, and Sound Res., Inc., San Diego, CA 92130, Ahmad.Abawi@HLSResearch.com)

The equivalent source technique is a method of computing scattering and radiation from a target by replacing the target by a distribution of discrete sources, whose complex amplitudes are determined by applying the boundary condition on the total field. The advantage of using this method, particularly in underwater acoustics, is that it essentially transforms the scattering and propagation problems into just a propagation problem, where the target can be treated as a distribution of sources rather than an impedance discontinuity. In this paper the equivalent source technique and different propagation models such as the normal mode, the parabolic equation and the fast field method are used to compute scattering from various target shapes/properties placed in the ocean waveguide. The results are compared with those obtained from other numerical codes. In addition, the relationship between the equivalent source technique with boundary element method (BEM) and the method of moments (MoM) is investigated.

10:30

5aUW8. Detecting small targets in strong bottom reverberation. Jinyun Ren and John S. Bird (School of Eng. Sci., Simon Fraser Univ., Burnaby, BC, Canada V5A 1S6, jren@cs.sfu.ca)

The detection of small targets moving or appearing in heavy bottom reverberation is a challenging problem for surveillance sonars. Conventional detection techniques, such as variable thresholding, are inefficient because they require that the target return be large enough to compete with the total reverberation return. Based on a new reverberation model, this paper proposes a target detection scheme that provides target subclutter visibility in the presence of reverberation. First it is shown through experimental evidence that bottom reverberation as seen by a stationary sonar is coherent or at least partially coherent from ping to ping. Therefore reverberation for a particular range cell can be modeled as a complex signal composed of a stationary or slowly varying mean plus a rapidly varying diffuse component. This mean is easily estimated using an adaptive mean

estimator and then removed so that the target need only compete with the diffuse component. Experimental results show that the detection gain over variable thresholding as measured by the coherent-to-diffuse ratio can be as much as 40 dB.

10:45

5aUW9. Channel effects on target characteristics in very shallow water. Kevin LePage (Naval Res. Lab., Code 7144, 4555 Overlook Ave. SW, Washington, DC 20375)

Very shallow water waveguides introduce channel effects which complicate the scattered response from objects. Here we model the effects of channel time spread and dispersion on the expected value of the time series backscattered by a sphere with and without uncertainty introduced by imperfect knowledge of the channel. We also model the predictable characteristics of the bottom reverberation against which detections are made. Results are intended to guide the design of dispersion and uncertainty-robust sonar systems in very shallow water. [Work supported by ONR.]

11:00

5aUW10. Mode extraction from acoustic backscatter in a waveguide. Shane Walker, Philippe Roux, and William A. Kuperman (Scripps Inst. of Oceanogr., Marine Physical Lab., 9500 Gilman Dr., La Jolla, CA 92093-0238, shane@physics.ucsd.edu)

Recently we introduced a technique for extracting the acoustic modes in a range-independent waveguide from measured data alone without *a priori* knowledge of the environment. This method requires a Vertical Line Array (VLA) to collect data from a moving broadband point source at a single depth in order to reconstruct the frequency-wavenumber (f - k) structure of the waveguide. Our goal is to apply this technique to the more complicated case of bottom reverberation using the VLA as both a source and receiver. As a broadband signal propagates away from the VLA, inhomogeneities along the water/bottom interface scatter some of the field back toward the VLA, mimicking a moving source. However, ambiguities inherent in assigning ranges to the received signal lead to ambiguities in the f - k structure. As a preliminary step, we remove these ambiguities by applying the technique to individual point scatterers at various known ranges along the bottom. These results lay a foundation from which to approach the problem of modeling the bottom as a distribution of point scatterers at unknown locations. Theory, simulation, and laboratory measurements are presented.

11:15

5aUW11. Nonlinear scattering of acoustical waves on sphere after passing the obstacle. Nickolay Zagrai, Andrey Zagrai, and Irene Starchenko (TSURE, 347928, Taganrog, GSP-17a, Nekrasovskiy, 44, Russia, znp@tsure.ru)

The problem of nonlinear interaction of acoustical waves in the system medium-obstacle-medium-sphere was considered. The expressions for secondary field of parametric array were obtained. The contributions of components were evaluated on separated processes of non-linear interaction before and after the obstacle as for the primary signals field as for the secondary fields of combinational frequencies waves. The possible practical applications were also discussed.

11:30

5aUW12. Separability of modal back-scattering matrix in shallow-water reverberation models. Jinrong Wu, Tianfu Gao, Erchang Shang, and Li Ma (Inst. of Acoust., Chinese Acad. of Sci., Beijing, China; No. 21, West Rd., Beisihuan, Beijing, China)

Normal mode model of the reverberation in shallow-water waveguides has been presented based on Born approximation. The key component of this model is the modal back-scattering matrix. Separability is one of the most important characteristics of the modal back-scattering matrix. The

separability or quasi-separability of this matrix could make it easy for modal back-scattering matrix inversion. In this paper, the modal back-scattering matrix due to the roughness of the bottom interface in shallow-water was obtained and its separability error was analyzed. Numerical simulation shows that the separability error increases with modal grazing

angle and density ratio of the sediment to the water, while decreases with sound speed ratio of the sediment to the water. It also shows that the effect of the separability error on shallow water reverberation level prediction is negligible. [Work supported by the National Science Foundation of China under Grant No 10474111.]

FRIDAY AFTERNOON, 20 MAY 2005

PLAZA C, 1:00 TO 2:50 P.M.

Session 5pAB

Animal Bioacoustics and ASA Committee on Standards: Behavioral Audiometric Methods in Animal Bioacoustics: The Search for Standards II

Edward J. Walsh, Cochair

Boys Town National Research Hospital, 555 N. 30th St., Omaha NE 68131

Ann E. Bowles, Cochair

Hubbs-Sea World Research Inst., 2595 Ingraham St., San Diego, CA 92109

Chair's Introduction—1:00

Contributed Papers

1:05

5pAB1. Dependence of detection threshold estimates in rabbit on method of constant stimulus parameters. Laurel H. Carney, Yan Gai (Inst. for Sensory Res., Dept. of Biomed. and Chemical Eng., Syracuse Univ., Syracuse, NY 13244), Kristina S. Abrams (Syracuse Univ., Syracuse, NY 13244), Fabio Idrobo (Boston Univ., Boston, MA 02215), and John Michael Harrison (Boston Univ., Boston, MA 02215)

The goal of this study was to estimate behavioral detection thresholds in quiet and noise for subsequent physiological studies of responses to near-threshold stimuli. The difficulty in measuring detection thresholds is maintaining stimulus control of behavior over schedule control. Tracking paradigms are efficient, but the large proportion of trials below threshold weakens stimulus control. The Method of Constant Stimuli controls the proportion of trials below, near, and above the SPL for a given percent correct. Here, percentages of trials at different SPLs and zero intensity catch trials were varied to determine proportions that yielded satisfactory stimulus control by a 500-Hz tone in quiet. Trials were initiated by nose-poke observing responses; a nose-poke reporting response within a 3-s window after tone onset resulted in food reinforcement. Reporting responses in the absence of tones resulted in timeouts. A metric for stimulus control was the ratio of the number of correct detections to the number of responses in the absence of the tone. This ratio was highest when approximately 20% of trials were near the 50%-correct SPL and 80% of trials were 20 dB higher. The ratio was lowest when significant proportions of catch and/or below-threshold trials were included. [Work supported by NIDCD DC-001641.]

1:20

5pAB2. Manatees masked thresholds for complex sounds. Edmund Gerstein, Laura Gerstein, Joseph Blue (Leviathan Legacy Inc., 1318 SW 14th St., Boca Raton, FL 33486), and Steve Forsythe (Naval Undersea Warfare Ctr., Newport, RI)

Acoustical masked thresholds of complex real world sounds were measured with West Indian manatees. Motor boat and species-specific calls were recorded, filtered and played back in a forced-choice two-alternative test. A method of constant stimuli was used to estimate the 50% detection thresholds of peak spectral components. Both white noise and recorded wild ambient spectra were used as continuous background noise for two sets of threshold estimates. Manatee calls were detected at and

within both white noise and wild ambient levels. Thresholds for broadband boat noise ranged from 9 to 29 dB above ambient conditions. Thresholds for idling boat noise were the highest. The levels were consistent with the subjects' auditory limits for the lower frequencies with the spectra. The manatee is well adapted to hear the species specific vocalizations under noisy conditions. The integration of multiple critical bands may result in loudness summation benefits which manatees can exploit at higher frequencies, however, at lower frequencies, near the manatees' limits of sensitivity, benefits of loudness summation were not apparent.

1:35

5pAB3. Testing marine mammal hearing with a vocal response paradigm and the method of free response. James J. Finneran, Donald A. Carder (U.S. Navy Marine Mammal Program, SPAWARSYSCEN San Diego, Code 2351, 53560 Hull St., San Diego, CA 92152, james.finneran@navy.mil), Carolyn E. Schlundt (EDO Professional Services, San Diego, CA 92110), and Sam H. Ridgway (Univ. of California, La Jolla, CA 92093-0612)

Behavioral hearing tests with marine mammals often use single interval experiments with a go/no-go paradigm, where the subject pushes a response paddle or similar device to indicate the presence of a hearing test tone. This approach is widely accepted, but the time required for the subject to physically move and/or interact with the response device and the requirement of one trial per reinforcement interval limit the speed at which data can be collected (typically 20–30 minutes for a threshold). An alternative to the single interval/paddle press approach is the use of a vocal response technique and the method of free response, where multiple trials are presented within a single reinforcement interval. An acoustic or vocal response eliminates the need for the subjects to physically move, allowing for faster responses and shorter inter-trial intervals. The method of free response allows multiple tone presentations between reinforcement periods, dramatically increasing the rate at which data may be collected, especially from diving marine mammals. This talk describes the application of these techniques to hearing tests of bottlenose dolphins, white whales, and California sea lions and the advantages and disadvantages compared to the single interval experiment. [Work supported by ONR.]

1:50

5pAB4. Noise localization ability in the Mongolian gerbil (*Meriones unguiculatus*). Srijata Sarkar, Kristina S. Abrams, and Laurel H. Carney (Inst. for Sensory Res. and Dept. of Biomed. and Chemical Eng., Syracuse Univ., Syracuse, NY 13244)

Understanding localization ability is essential for comparison to physiological studies of binaural neurons. Heffner and Heffner [Beh. Neuro. **102**, 422 (1998)] reported relatively poor ability of the gerbil to localize a 100 ms noise burst (approximately 75% correct at 27° separation.) Their experimental setup required the gerbil to enter an observing response compartment, initiating a noise burst from one of two speakers separated by the test angle. The animal was required to respond by entering a reporting compartment, positioned 90° to the right or left, regardless of speaker location. This required mapping from speaker location to response location. In this study, response mapping was avoided in an effort to improve performance. After an observing response (jumping on a platform in the center of a circular cage), the animal responded by moving directly towards the active speaker. The results of this study were consistent with those of Heffner and Heffner at small angles of separation. [For 180° separation, this task resulted in poorer performance than the previous study, presumably due to lack of control of head orientation during stimulation, leading to front-back confusions.] In conclusion, the poor localiza-

tion ability of gerbils is apparently not explained by response mapping. [Work supported by NIDCD R01-DC001641.]

2:05

5pAB5. Strain differences in auditory sensitivity in canaries bred for song and plumage. Amanda Lauer, Elizabeth Brittan-Powell, Bernard Lohr, and Robert Dooling (Dept. of Psych., Univ. of Maryland, College Park, MD 20742)

Canaries have been domesticated for well over 100 years. They have been bred for specific characteristics such as song, plumage, or body shape. Here we measured audiograms in several canary strains using both behavioral (operant conditioning) and physiological (auditory brainstem responses) methods. Overall, there was a good correspondence between behavioral and ABR audiograms, but there were differences in audiogram shape and sensitivity between strains. Belgian Waterslager canaries bred for low-pitched song had elevated high frequency thresholds, while Spanish Timbrado canaries bred for high-pitched song had slightly better high frequency thresholds than canaries that were not bred for song. Other strains showed intermediate sensitivity. These results suggest there can be more variability in hearing within a species than previously thought and that these differences are correlated with other behavioral or morphological traits. [Work supported by DC00198, DC001372, DC04664-01A2, and DC005450.]

2:20–2:50

Panel Discussion

Panelists: Robert Dooling, Richard Fay, Henry Heffner, Paul Nachtigall, Constantine Trahiotis, William Yost.

FRIDAY AFTERNOON, 20 MAY 2005

PLAZA A, 2:00 TO 4:15 P.M.

Session 5pBB

Biomedical Ultrasound/Bioresponse to Vibration and Physical Acoustics: Audible-Frequency Medical Diagnostic Methods Including Multimode Techniques II

Hans Pasterkamp, Cochair

Pediatrics Dept., Univ. of Manitoba, 840 Sherbrooke St., Winnipeg, MB R3A 1R9, Canada

Thomas J. Royston, Cochair

Dept. of Mechanical Engineering, Univ. of Illinois at Chicago, 842 W. Taylor St., Chicago, IL 60607-7022

Chair's Introduction—2:00

Invited Papers

2:05

5pBB1. Monitoring changes of human airways by acoustical means. Hans Pasterkamp (Dept. of Pediatrics, Univ. of Manitoba, CS5120-840 Sherbrook St., Winnipeg, MB, Canada R3A 1S1, pasterkamp@umanitoba.ca)

The state of human airways is typically assessed by spirometry, i.e., by measuring airflow during maximum forced expiration. The required effort excludes subjects who cannot or will not perform such maneuvers. Turbulent airflow during normal breathing generates sound that can be recorded at the chest surface. Changes in airway caliber, e.g., by constriction or dilation, and concurrent changes in the tension of the airway walls affect the relation of airflow and breath sounds. This relation is complex and incompletely understood. However, the possibility to assess changes in the state of human airways by the relation of airflow and breath sounds promises an opportunity to develop auxiliary methods to spirometry. Two regions of the human airways have attracted the interest of researchers in respiratory acoustics. Breath sounds over the trachea (windpipe) increase in their airflow-specific intensity on narrowing of the upper airways, e.g., due to abnormal anatomy, infectious causes etc. Furthermore, spectral characteristics of tracheal sounds may indicate the region of abnormality. On the other hand, breath sounds over the chest change in their airflow-specific intensity with constriction of airways in the lung, e.g., due to asthma. It may therefore be possible to monitor the effectiveness of treatment by acoustical means.

5pBB2. A system to evaluate and compare transducers used for lung sound research. Steve S. Kraman, George R. Wodicka, Gary A. Pressler, and Hans Pasterkamp (Kentucky Clinic J515, Univ. of Kentucky, Lexington, KY 40536) (Purdue Univ., West Lafayette, IN 47907) (Purdue Univ., West Lafayette, IN 47907) (Univ. of Manitoba, Winnipeg, MB, Canada)

Since lung sounds were first recorded and analyzed in the 1960s, a variety of custom made or adapted transducers have been used for detecting these sounds. There is no standard lung sound transducer nor have those in use been adequately compared. To address this problem, a test platform was constructed that provides a stable and reproducible sound source coupled to a viscoelastic surface that is mechanically and acoustically similar to human skin and subcutaneous tissue as judged from comparisons with similar thicknesses of fresh meat and fat. The device was designed to be equally suitable for air-coupled microphones and for accelerometers. When driven by a broadband noise source, the acoustic amplitude at the surface of the platform was found to be relatively uniform from 100 to 1200 Hz. The system was used to evaluate a variety of lung sound transducers. They were found to vary significantly in frequency response with no device appearing to be adequate as an all-purpose lung sound sensor. It is concluded that lung sound characteristics are likely to be significantly affected by the type of sensor used to acquire the sound. This makes comparisons between differently equipped laboratories difficult.

5pBB3. Vascular sounds as an indicator of hemodialysis access patency. Hansen Mansy, Silas Hoxie, Nilesh Patel, and Richard Sandler (Rush Medical College, Chicago, IL 60607)

Background: Vascular access for renal dialysis is a lifeline for about 200 000 individuals in North America. Stethoscope auscultation of vascular sounds has some utility in the assessment of access patency, yet may be highly skill dependent. The objective of this study is to identify acoustic parameters that are related to changes in vascular access patency. Methods: Fifteen patients participated in the study. Their vascular sounds were recorded before and after angiography, which was accompanied by angioplasty in most patients. The sounds were acquired using two electronic stethoscopes and then digitized and analyzed on a personal computer. Results: Vessel stenosis changes were found to be associated with changes in acoustic amplitude and/or spectral energy distribution. Certain acoustic parameters correlated well (correlation coefficient=0.98, p less than 0.0001) with the change in the degree of stenosis, suggesting that stenosis severity may be predictable from these parameters. Parameters also appeared to be sensitive to modest diameter changes (greater than 20%), (p less than 0.005, Wilcoxon rank sum test). Conclusions: Computerized analysis of vascular sounds may have utility in vessel patency surveillance. Further testing using longitudinal studies may be warranted. [Work supported by NIH/NIDDK DK 59685.]

3:20–3:30 Break

Contributed Papers

3:30

5pBB4. Computational and experimental models for sound transmission in the pulmonary system and chest to aid in medical diagnosis. Serhan Acikgoz, Thomas J. Royston, M. Bulent Ozer (Univ. of Illinois at Chicago, 842 W. Taylor St. MC 251, Chicago, IL 60607), Hansen A. Mansy, and Richard H. Sandler (Rush Univ. Medical Ctr., Chicago, IL 60612)

Acoustic wave propagation in the pulmonary system and torso is simulated by coupling a numerical acoustic boundary element model that predicts sound propagation throughout the solid tissues with a proven comprehensive analytical model for sound wave propagation in the airways of the bronchial tree that is valid up to at least 2 kHz. This approach enables modeling various pathologies that result in structural changes in the lung and/or changes in breath sound source and strength. The model may be also used to predict the resulting acoustic changes measured by acoustic sensors, e.g., stethoscopes, accelerometers, or other skin-contact sensors. Experimental studies in a novel lung phantom model are used to partially validate the computational model. This study focuses on low audible frequencies, i.e., less than 2 kHz. This range encompasses naturally generated respiratory sounds that have been shown to have diagnostic value, as well as externally-introduced vibro-acoustic stimuli used for diagnosis. [Work supported by NIH EB 003286-01.]

3:45

5pBB5. Recent developments in the characterization of anisotropic, biological media utilizing magnetic resonance elastography. Anthony J. Romano, Nicolas P. Valdivia, Philip B. Abraham, Joseph A. Bucaro (Naval Res. Lab., 4555 Overlook Ave, Washington, DC 20375-5320, romano@pa.nrl.navy.mil), Philip J. Rossman, and Richard L. Ehman (Mayo Clinic and Foundation, Rochester, MN 55905)

In this paper, we present recent developments in the characterization of anisotropic, biological media utilizing displacements measured with dynamic MRE. In one approach, a waveguide constrained analysis is applied to physical structures (such as muscle), which evaluates the velocities of wave propagation along arbitrarily oriented fibers or fiber bundles in a local fashion, utilizing a spectral filter and a sliding window, spatial Fourier transform. In a second approach, the anisotropic equations of elasticity, in variational form, are inverted for a determination of the complex moduli comprising the medium. Issues concerning ill-conditioning, the effects of boundary conditions, and multiaspect excitation requirements will be addressed, as well as a discussion concerning the models which are appropriate for various physical situations. These two complementary methods of analysis will be demonstrated utilizing both phantom as well as *in vivo* dynamic displacement measurements.

4:00

5pBB6. Local inversion of transient shear-wave propagation for elasticity and viscosity mapping in soft tissues. Jeremy Bercoff, Mickael Tanter, and Mathias Fink (Laboratoire Ondes et Acoustique, ESPCI, 10 rue Vauquelin, 75005 Paris, France)

Observation of transient shear-wave propagation in soft tissue is of great interest for the study of tissue viscoelastic properties. In previous work, we introduced a technique, called Supersonic Shear Imaging (SSI), able to generate transient shear waves using the acoustic radiation force and image their propagation in real time in soft tissues. In this work, a

local inverse problem of the shear wave propagation permitting the recovery of shear elasticity and viscosity is presented. Compressional and shear waves are decoupled by applying the curl operator to the experimental 2-D displacement field. The role of viscosity on the accuracy of the elasticity estimation is studied. The influence of out of plane shear propagation on the inversion algorithm is discussed. Finally, in media presenting shear viscoelasticity heterogeneities, finite difference simulations are used to study the spatial resolution of the algorithm and its sensitivity to the signal-to-noise ratio. Experiments on calibrated tissue-mimicking phantoms presenting different viscoelastic properties are presented validating the simulation results. First *in vivo* results on female breasts are presented.

FRIDAY AFTERNOON, 20 MAY 2005

REGENCY F, 1:30 TO 4:45 P.M.

Session 5pNS

Noise and Psychological and Physiological Acoustics: Workshop on Methods for Community Noise and Annoyance Evaluation II

Brigitte Schulte-Fortkamp, Chair

Technical Univ. Berlin, Inst. of Technical Acoustics, Secr TA 7, Einsteinufer 25, Berlin 10587, Germany

Invited Papers

1:30

5pNS1. Auditorium Mundi soundscape in the narration of the world. Artistic and scientific aspects in the presentation of soundscapes. Harald Brandt (Philippsbergstrasse 40, D-65195 Wiesbaden, Germany, haraldbrandt@yahoo.fr)

Auditorium Mundi is a project for the exploration and presentation of sounds and soundscapes of the whole world. The aim is an installation of a museum of audible perception in interrelation with other senses. The concept of the museum is an acoustic planetarium in which the locations of the world are put in place of the stars. Also located here are world archives for sound in which the acoustic and musical multitude of this planet will be preserved. The archives of sound are available world-wide through the internet. The museum is, at the same time, a laboratory for the development of new formats in scientific definitions and the multimedia presentation of acoustic phenomena. It is a show-place for artistic programs and an auditorium where questions from the area of acoustic as well as new developments of industry will be accessible to a wider public. A place of innovation, where concepts for other museums, promoters and scientific institutions can be developed. The leitmotif of Auditorium Mundi is the question how man influences the soundscapes in which he lives and how strongly he himself is influenced by the noise, the resonance and the dissonance between various ways of life.

1:50

5pNS2. An L_{Aeq} is not an L_{Aeq} . Dick Botteldooren, Tom De Muer, Bert De Coensel (Acoust. Group, Dept. of Information Technol., Ghent Univ., St. Pietersnieuwstraat 41, B-9000 Ghent, Belgium), Birgitta Berglund (Stockholm Univ., SE-106 91 Stockholm, Sweden), and Peter Lercher (Univ. of Innsbruck, A-6020 Innsbruck, Austria)

Classical dose response relationships for environmental noise annoyance have been based on L_{dn} or L_{den} . These exposure measures are essentially based on an energy averaging measure, L_{Aeq} . Differences between groups of sources (e.g., continuous or event based) are accounted for by using separate dose-effect relationships. In society today, one often sees that event loudness is traded for number of events which is perfectly acceptable within the L_{Aeq} based annoyance concept. Clearly a more unified theory for noise annoyance is needed to fully account for the effect of such trade-offs. In this paper a model implementing such a theory is presented. The perceptual model starts from the premises that a sound event has to be noticed for it to contribute to overall annoyance. The model accounts for the fact that noticing a noise event not only depends on the level of the event itself but also on background noise, sound insulation and acoustic characteristics of the dwelling, level of attention, etc., the severity of the effect of a noticed sound on overall annoyance is assumed to primarily depend on the signal to noise ratio. The model allows to account for modifiers such as previous exposure, noise sensitivity, and coping. The model results are compared to the findings of a recent field experiment. Conclusions based on calculated and experimental trends will be presented.

FRIDAY AFTERNOON, 20 MAY 2005

REGENCY D, 1:00 TO 5:00 P.M.

Session 5pSC

Speech Communication: Speech Production and Perception II (Poster Session)

Bryan W. Gick, Chair

*Dept. of Linguistics, Univ. of British Columbia, 1866 Main Mall E 270, Vancouver, BC V6T 1Z1, Canada**Contributed Papers*

All posters will be on display from 1:00 p.m. to 5:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:00 p.m. to 3:00 p.m. and contributors of even-numbered papers will be at their posters from 3:00 p.m. to 5:00 p.m.

5pSC1. Perceptual development of phonotactic features in Japanese infants. Ryoko Mugitani (NTT Commun. Sci. Labs., NTT Corp., 2-4 Hikari-dai, Seika-cho, Soraku-gun, Kyoto, 619-0237, Japan), Laurel Fais (Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z4), Sachiyo Kajikawa, Shigeaki Amano (NTT Corp., Seika-cho, Soraku-gun, Kyoto, 619-0237, Japan), and Janet Werker (Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z4)

Acceptable phonotactics differ among languages. Japanese does not allow consonant clusters except in special contexts and this phonotactic constraint has a strong effect on adults speech perception system. When two consonants follow one another in nonsense words, adult Japanese listeners hear illusory epenthetic vowels between the consonants. The current study is aimed at investigating the influence of language-specific phonotactic rules on infants' speech perception development. Six-, 12-, and 18-month-old infants were tested on their sensitivity to phonotactic changes in words using a habituation-switch paradigm. The stimuli were three nonsense words: "keet (/ki:t/)," "keets (/ki:ts/)," and "keetsu (/ki:tsu/)." "Keetsu" perfectly follows Japanese phonotactic rules. "Keets" is also possible in devoicing contexts in fluent speech, but the acceptability of "keets" for adult native Japanese speakers is much less than "keetsu." "Keet" is phonotactically impossible as a Japanese word. The results indicate the existence of a developmental change. Twelve months and older infants detected a change from the acceptable Japanese word "keetsu" to the possible but less acceptable word "keets." However, discrimination between "keets" and the non-Japanese "keet" is seen only in 18-month infants. Implications for infants' speech perception with relation to language specific phonotactical regularities will be discussed.

5pSC2. Temporal evidence against the production of word-internal syllable structure. Melissa A. Redford (Dept. of Linguist., 1290 Univ. of Oregon, Eugene, OR 97403, redford@darkwing.uoregon.edu)

A production study examined whether the temporal pattern that cues onset cluster syllabification in English should be attributed to syllable structure per se or to coarticulatory processes that depend on segment identity and sequencing, but not on syllable structure. Experiment 1 tested for independent effects of syllable structure on intervocalic stop-liquid consonant duration. Three speakers produced stimuli in which sequence order, boundary strength, and stress were manipulated. The results showed that stimuli with word-internal sequences were produced with similar stop-liquid duration ($S:L$) ratios regardless of sequence order or stress.

When a word boundary split the consonants, $S:L$ ratios were much smaller, primarily because liquids were longer. Experiment 2 was aimed at distinguishing between alternate explanations for the pattern observed in Experiment 1. Data from three Russian speakers suggested that the English pattern was due to word-internal liquid reduction rather than to boundary lengthening. The $S:L$ ratios in Russian were not affected by sequence order or boundary strength. Further, the ratios were significantly different from the word-internal $S:L$ ratios of English, but not from the word-peripheral ones. Overall, the results argue against a causal relationship between syllable structure and articulatory timing, and for word-sized articulatory units in English.

5pSC3. Perception of coarticulated speech with contrastively enhanced spectrotemporal patterns. Travis Wade and Lori Holt (Dept. Psych., Carnegie Mellon Univ., 5000 Forbes Ave., Pittsburgh, PA 15213, twade@andrew.cmu.edu)

High-level contrastive mechanisms cause perception of auditory events to be influenced by spectral and temporal properties of surrounding acoustic context, and may play a role in perceptual compensation for coarticulation in human speech. However, it is unknown whether auditory contrast is incorporated optimally to compensate for different speakers, languages and situations or whether amplification of the processes involved would provide additional benefit, for example, in the perception of hypoarticulated speech, under adverse listening conditions, or in an incompletely acquired language. This study examines effects of artificial contrastive modification of spectrotemporal trajectories on the intelligibility of connected speech in noise by native and non-native listeners. Adopting methods known to improve automatic classification of speech sounds, we model contrast-providing context as an averaged estimated vocal tract function (LPC-derived log area ratio coefficient vector) over a Gaussian-weighted temporal window. Local coefficient values are adjusted from this context based on previously observed contrastive perceptual tendencies, and the intelligibility of the resulting speech is compared with that of unmodified trajectories across listener language backgrounds. Results are discussed with respect to implementation and applicability of general auditory processes.

5pSC4. Lip and jaw closing gesture durations in syllable final voiced and voiceless stops. Willis Warren (Univ. of Texas Dept of Linguist., Calhoun 501, 1 University Station B5100, Austin, TX 78712-0198, warisill@mail.utexas.edu) and Adam Jacks (Univ. of Texas, Austin, TX 78712)

It is generally assumed that the longer vowel durations in voiced final consonant contexts can be attributed to longer vowel steady state durations wherein the jaw remains stable for a longer period of time (see diagrams in Chen, 1970; Summers, 1987). Chen (1970) argued that the rate of closure duration was the best explanation for the vowel length effect. However, the closure durational differences he reported were very small relative to the reported differences in vowel duration with an average of 17 ms to 92 ms. Lip and jaw movements for two speakers were measured and synchronized with acoustic data in the CSD department at the University of Texas using the head-mounted lip-jaw movement transducer (HML-JMT) of Barlow *et al.* (1983). This study provides much stronger evidence for Chen's hypothesis, suggesting that there is a more dynamic arc to the jaw. While Chen (1970) gave data showing the closing gesture accounted for less than 20% of the overall difference, the data given here, shows that the duration of the closing gesture can account for 80% of the total vowel durational differences, and the remainder can be accounted for in the opening gesture and the relative timing of vocal cord cessation.

5pSC5. Assimilation context effects as cue grouping phenomena. Nicole Couture, Sarah Godbois (Dept. of Psych., Salem State College, 352 Lafayette St., Salem, MA 09170, gow@helix.mgh.harvard.edu), and David W. Gow, Jr. (Massachusetts General Hospital, Boston, MA 02114)

Listeners show bidirectional context effects in the perception of spontaneously assimilated speech. One account of these phenomena, feature cue parsing theory, suggests that both regressive and progressive context effects result from the same perceptual grouping process of aligning feature cues with assimilated word offsets and assimilating word onsets. This hypothesis was examined through two form priming experiments that explored the interpretation of spontaneously assimilated items under manipulations that either favored or disfavored the perceptual grouping of feature cues associated with assimilated and assimilating items. In both experiments, listeners heard phrases such as right berries in which a word final coronal segment (*/t/*) assimilated the non-coronal place of a subsequent word onset (labial */b/*) to form an item that sounded like one of its lexical competitors (ripe). In Experiment 1 spatial continuity was manipulated through the use of same and different side lateralized presentation of assimilated words and their contexts. In Experiment 2 continuity in speaker identity was manipulated through cross-splicing. Analyses contrasted the pattern of priming found in continuous versus non-continuous contexts. The results are discussed in the terms of competing interpretations of assimilation context effects. [Work supported by the NIH.]

5pSC6. Relationships among perceived sexual orientation, perceived height, and perceived speech clarity. Benjamin Munson and Elizabeth C. McDonald (Dept. Speech-Lang.-Hearing Sci., Univ. of Minnesota, 115 Shevlin Hall, 164 Pillsbury Dr., SE, Minneapolis, MN 55455 Munso005@umn.edu)

Previous research [B. Munson, *Linguist. Soc. Am.* (2005)] showed that listeners rate self-identified gay/lesbian/bisexual (GLB) talkers as more-GLB sounding than self-identified heterosexual talkers when listening to readings of single words in isolation. In this investigation, we examine relationships among measures of perceived sexual orientation (PSO) and two other perceptual measures made from the same talkers' speech, perceived height and perceived speech clarity. Ratings were collected from three independent groups of listeners. Regression analyses showed that approximately 80% of the variance in PSO of women's voices was predicted by measures of perceived clarity (women rated as speaking more clearly were more likely to be rated as heterosexual-sounding than women rated as speaking less clearly) and perceived height (women rated as sounding tall were more likely to be rated as GLB-sounding than women

rated to sound short). In contrast, 50% of the variance in men's sexual orientation was predicted by measures of perceived clarity (men rated as speaking more clearly were more likely to be rated as GLB-sounding than men rated as speaking less clearly). These results suggest that ratings of PSO from speech may be mediated by judgments of other parameters that are robustly coded in short samples of speech.

5pSC7. Relative contributions of feedback and editing in language production: Behavioral & articulatory evidence. Corey T. McMillan, Martin Corley (Univ. of Edinburgh, PPLS, 7 George Square, Rm. F23, Edinburgh, EH8 9JZ, UK, Corey.McMillan@ed.ac.uk), Robin J. Lickley (Queen Margaret Univ. College, Edinburgh, EH12 8TS, UK), and Robert J. Hartsuiker (Univ. of Ghent, 9000 Ghent, Belgium)

Psychologists normally attribute the surfacing of phonological speech errors to one of two factors: editing of the speech plan [Levelt (1989)] or feedback between word and phoneme levels [Dell (1986)]. This paper assesses the relative contributions of each factor, focusing on the perception and articulation of elicited speech errors. Experiments one and two measure the likelihood of phonological exchange errors as a function of phonetic similarity [Frisch (1996)], using the SLIP paradigm and a tongue-twister task. Both experiments show that error likelihood increases with phonetic similarity between intended and actual utterance, an effect easy to account for in terms of feedback but not in terms of editing. Experiment three uses EPG to analyze the tongue-twister utterances: many errors occur at the articulatory level but are not easily perceived in the speech signal. Preliminary analysis suggests three patterns of error: (1) substitution of segments, which may be the result of editing; (2) simultaneous double articulation, hypothesized to be the result of residual activation due to feedback; and (3) overlapping double articulation, representing partial execution of one articulation before substitution with another. Taking these findings together, we hope to evaluate the relative contributions of editing and feedback to phonological speech errors.

5pSC8. Acoustical study of the development of stop consonants in children. Annika K. Imbrie (Speech Commun. Group, MIT, Rm. 36-545, 77 Massachusetts Ave., Cambridge, MA 02139, imbrie@mit.edu)

This study focuses on the acoustic patterns of stop consonants and adjacent vowels as they develop in young children (ages 2;6-3;3) over a six month period. Over forty different acoustic measurements were made on each utterance, including durational, amplitude, spectral, formant, and harmonic measurements. These acoustic data are interpreted in terms of the supraglottal, laryngeal, and respiratory actions that give rise to them. Data show that some details of the child's gestures are still far from achieving the adult pattern. Significant findings include a high number of bursts in children's stop productions, possibly due to greater compliance of the active articulator. The incidence of multiple bursts decreases over the period of the study. Measurements of the burst spectra show that the production of all three places of articulation is close to normal. Finally, children's voiced stop productions display a delayed onset of voicing for all three places of articulation, indicating difficulty with coordination of closure of the glottis and release of the primary articulator. This measurement is found to decrease toward adult values over six months. Analysis of longitudinal data on young children will result in better models of the development of motor speech production. [Work supported by NIH grant DC00075.]

5pSC9. Relations between speech sensorimotor adaptation and perceptual acuity. Virgilio Villacorta, Joseph Perkell, and Frank Guenther (Res. Lab of Electron., Mass. Inst. of Tech., 50 Vassar St., Ste. 36-511, Cambridge, MA 02139)

The goal of this research is to study the auditory component of feedback control in speech production. This experiment investigates auditory sensorimotor adaptation (SA) as it relates to speech production: the pro-

cess by which speakers alter their speech production in order to compensate for perturbations of normal auditory feedback. Specifically, the first formant frequency (F_1) was shifted in the auditory feedback heard by naive adult subjects as they produced vowels in single syllable words. Results presented previously indicate that the subjects demonstrate compensatory formant shifts in their speech. The current study investigates the relationship between compensatory adaptation and speaker perceptual acuity. Subjects from the SA study were tested for their ability to discriminate vowel tokens differing in F_1 frequency. Preliminary results indicate that subjects with greater perceptual acuity also demonstrated greater ability to adapt, with a significant cross-subject correlation (p less than 0.05). The relation between these experimental findings and DIVA, a neurocomputational model of speech motor planning by the brain, will also be discussed. [Work supported by NIDCD Grant R01-DC01925.]

5pSC10. Glottal-wave and vocal-tract-area-function estimations from vowel sounds based on realistic assumptions and models. Huiqun Deng, Rabab K. Ward, Micheal P. Beddoes (Elec. and Comput. Eng. Dept., Univ. of British Columbia, Canada), Murray Hodgson, and Bryan Gick (Univ. of British Columbia, Canada)

Estimating glottal waves by inverse filtering vowel sounds and deriving vocal-tract area functions (VTAFs) from vocal-tract filter (VTF) estimates require that VTF models be realistic and that VTF estimates contain no effects of open glottises and glottal waves. In this study, VTFs are modeled to have lip reflection coefficients with low-pass frequency responses; to minimize the effects of open glottises and glottal waves on the estimates, VTFs are estimated from sustained vowel sounds over closed glottal phases, assuming that the glottal waves are periodically stationary random processes. Since incomplete glottal closures are common, VTF estimates may contain the effects of glottal loss. To eliminate the effects of glottal loss in the VTF estimates, lip-opening areas must be known. Theoretically, estimates of glottal waves and VTAFs corresponding to large-lip-opening vowel sounds are less affected by the glottal loss than those corresponding to small-lip-opening vowel sounds. The VTAFs and glottal waves estimated from vowel sounds produced by several subjects are presented. The normalized VTAFs estimated from large-lip-opening sounds are similar to that measured from an unknown subjects magnetic resonance image. Over closed glottal phases, the glottal waves are non-zero. They increase during vocal-fold colliding, and decrease or even increase during vocal-fold parting.

5pSC11. Cross-modal perception of vowels, stops, and approximants using reduced-formant stimuli. Philip Rubin (Haskins Labs., 270 Crown St., New Haven, CT 06511-6204 and Yale Univ., School of Medicine, New Haven, CT, rubin@haskins.yale.edu), Catherine Best (Haskins Labs., New Haven, CT 06511-6204), Gordon Ramsay, Stephen Frost, and Bruno Repp (Haskins Labs., New Haven, CT 06511-6204)

This paper explores the relationship between audio and video representations of speech, by examining how information in particular formants is complemented by facial information, for different classes of speech sounds. Specifically, we investigate whether acoustic information removed by deleting formant combinations that usually signal contrasts between vowels (F_1, F_2), stops (F_2) and approximants (F_3) can be replaced by optical information derived from the talking face. Simultaneous audio and video recordings were made of speakers uttering CVC nonsense syllables constructed from vowels, stops, and approximants exhibiting varying degrees of facial motion and varying dependence on individual formant cues. Synthetic stimuli were then resynthesized from combinations of one, two, and three formants using a parallel formant synthesizer. Subjects performed an auditory identification task with stimuli presented in an audio-only condition, followed by a separate block of stimuli presented in an audio-visual condition. Source effects were examined using an inverse-filtered source and a synthetic source with constant source characteristics. There were no differences between source conditions, but differences were obtained between the audio and audio+video conditions that reflect the

relationship between facial and formant dynamics. These results complement recent research on perception of reduced video information and complete audio information [L. Lachs and D. B. Pisoni, *J. Acoust. Soc. Am.* **116**, 507–518 (2004)]. [Work supported by NIH.]

5pSC12. Analysis of vowel and speaker dependencies of source harmonic magnitudes in consonant-vowel utterances. Markus Iseli, Yen-Liang Shue, and Abeer Alwan (UCLA, 405 Hilgard Ave., Los Angeles, CA 90095, iseli@ee.ucla.edu)

It is assumed that voice quality characteristics are mainly manifested in the glottal excitation signal. As [Holmberg *et al.*, *J. Speech Hear. Res.* **38**, 1212–1223 (1995)] showed, there is a correlation between low-frequency harmonic magnitudes of the glottal source spectrum and voice quality parameters. In this study, we assess the influence of vowel and speaker differences on the difference between the first and the second harmonic magnitudes, $H_1 - H_2$. The improved harmonics correction formula introduced in [Iseli *et al.*, *Proceedings of ICASSP*, Vol. 1 (2004), pp. 669–672] is used to estimate source harmonic magnitudes. $H_1 - H_2$ is estimated for consonant-vowel utterances where the vowel is one of the three vowels /a, i, u/ and the consonant is one of the six plosives in American English /b, p, d, t, g, k/. Several repetitions of each of the utterances, spoken by two male and two female talkers, are analyzed. Other measurements, such as fundamental frequency, F_0 , and energy are estimated pitch synchronously. Results are compared to current literature findings. [Work supported by NSF.]

5pSC13. Pharynx depth and tongue height in relationship to intrinsic F_0 of vowels. D. H. Whalen, Leonardo Oliveira, and Bryan Gick (Haskins Labs, 270 Crown St., New Haven, CT 06511, whalen+AEA@haskins.yale.edu)

Intrinsic F_0 (IF_0) is the tendency for high vowels to have higher F_0 s than low vowels. It is universal and apparently automatic, but the exact mechanism is still unclear. Here, ultrasound measurements of the tongue were used to determine whether vowel height is the best predictor of IF_0 . Ten native speakers of English spoke keywords for the eleven non-diphthongal vowels while the lips, jaw, and tongue were imaged with the HOCUS ultrasound/Optotrak system [D. H. Whalen *et al.*, *J. Sp. Lang. Hear. Res.* (in press)], which tracks the head and lips with optical markers while the tongue is imaged with ultrasound. The tongue edge can be extracted and rotated relative to the head, indicating changes in tongue height and pharyngeal depth. Preliminary analysis indicates that IF_0 is better correlated with the most anterior point on the tongue root (indicating pharyngeal depth) than with the highest point on the tongue. It thus seems likely that the shape of the pharynx is more of a determinant than tongue height for IF_0 . Because pharynx depth and vowel height covary in most languages (including English), the earlier description is somewhat accurate; IF_0 may be more accurately described as correlated to the tongue root.

5pSC14. Acoustically-guided articulation patterns for vowel production. Brad Story (Dept. of Speech, Lang., and Hearing Sci., Univ. of Arizona, P.O. Box 210071, Tucson, AZ 85721, bstory@u.arizona.edu)

The purpose of this study was to investigate the link between the vocal tract deformation patterns obtained from statistical analyses of articulatory data (i.e., measured area functions) and the acoustic properties of a realistic neutral vocal tract shape. The starting point was a previously reported statistical analysis (i.e., PCA) of a speaker-specific set of measured vowel area functions. It resulted in a mean area function, and two basis functions that perturb the vocal tract into vowel-like shapes. In the present study, acoustic sensitivity functions were calculated for the same mean area function. Then, four separate sum and difference combinations of the sensitivity functions were used directly to perturb the mean area function. The

results showed that these sum and difference combinations created deformation patterns with nearly the same variation along the vocal tract length as the basis functions derived from articulatory data. The procedure was then applied to other speaker-specific sets of area functions and the results were similar. [Work supported by NIH R01-DC04789.]

5pSC15. The effect of spatial frequency information in central and peripheral vision on natural gaze patterns and audiovisual speech perception. Julie N. Buchan, Amanda Wilson, Martin Paré, and Kevin G. Munhall (Queen's Univ., Kingston, ON, Canada, 2jnb@qmlink.queensu.ca)

It has been known for some time that visual information plays an important role in how we perceive speech. The present research examines how this visual information is processed in central and peripheral vision. The perception of visual images is carried out in the nervous system by a set of spatial frequency-tuned channels and the sensitivity to spatial resolution varies across the retina. The experiment was conducted using a gaze-contingent display system that records a person's eye position and then displays different information to their peripheral and central vision. The amount of high spatial frequency information presented in the central and peripheral visual fields was manipulated and the effect on both audiovisual speech perception and natural gaze patterns was measured. Preliminary results show that when peripheral information is no longer adequate, gaze patterns become altered to gather important visual information. This suggests that information across the whole visual field influences audiovisual speech behavior.

5pSC16. Perception of a Japanese moraic obstruent in an identification task. Shigeaki Amano, Ryoko Mugitani, and Tessei Kobayashi (NTT Commun. Sci. Labs., NTT Corp., 2-4 Hikari-dai, Seika-cho, Souraku-gun, Kyoto 6190237, Japan)

Native Japanese speakers perceive a moraic obstruent when the closure duration between successive moras is longer than a normal obstruent. However, it has not been well clarified how the perception of the moraic obstruent relates to the closure and neighboring moras' duration. To investigate this point, a perceptual experiment was conducted with stimuli consisting of two-mora nonsense syllables with a closure between the moras (/bipa/, /guku/, /kuku/, /kuto/, and /tapi/). The closure duration was shortened or lengthened in 10-ms steps. The duration of the first mora was modified by changing its vowel duration to 50%, 100%, and 150% by the STRAIGHT method. Forty native Japanese speakers identified whether the nonsense syllable contained a moraic obstruent with the 2AFC method. Results showed that the perceptual boundaries of the moraic obstruent were 182 ms (SD=13.2 ms), 204 ms (SD=14.7 ms), and 222 ms (SD=16.6 ms), respectively, with 50%, 100%, and 150% vowel durations in the first mora, and that a logistic function fits the identification ratio very well. The results indicated that the perception of a moraic obstruent depends on both closure duration and the duration of neighboring moras. They also suggest that the perception would be categorical.

5pSC17. Nasals and nasalization: The interplay between segmental duration and coarticulation. Samantha E. Sefton and Patrice S. Beddor (Dept. of Linguist., Univ. of Michigan, 4080 Frieze Bldg., Ann Arbor, MI 48109, beddor@umich.edu)

Two experiments were conducted on English to test the hypothesis that the temporal extent of coarticulatory nasalization is inversely related to the duration of the source of coarticulation. Experiment 1 investigated patterns of carryover (NVC) and anticipatory (CVN) vowel nasalization, and the overall timing of nasalization across the syllable. Acoustic analysis of CVN and NVC words (e.g., seen/niece, pin/nip, tame/mate, ten/net) produced by five speakers showed that, although vowel nasalization is temporally more extensive in CVN than in NVC sequences, *N* is shorter in CVN than in NVC, so that there was no significant overall difference in

total nasalization between the two syllable types. Experiment 2 tested similar timing patterns in CVNC syllables, where *N* duration varied as a function of the voicing of the following *C* (e.g., wince/wins, sent/send, can't/canned). Acoustic measures of productions of five English speakers again showed that the shorter the *N*, the temporally more extensive the vowel nasalization. Findings for selected other languages (Thai, Italian) provide further evidence of the inverse relation. These findings have implications for theories of coarticulation and theories of phonological change. [Work supported by NSF.]

5pSC18. Perception of a Japanese moraic obstruent in a discrimination task. Tessei Kobayashi, Ryoko Mugitani, and Shigeaki Amano (NTT Commun. Sci. Labs., NTT Corp., 2-4, Hikaridai, Seika-cho, Soraku-gun, Kyoto, Japan 619-0237)

Previous research with an identification task suggests that native Japanese speakers perceive a moraic obstruent categorically (Amano et al., 2005). However, it is still unclear whether the suggestion would be confirmed by other tasks. This study used an AX discrimination task to investigate whether the performance is highly sensitive around the perceptual boundaries that were obtained from the identification task. Native Japanese speakers (*N*=40) were presented with a pair of two-mora nonsense syllables (/bipa/, /guku/, /kuku/, /kuto/) and then required to judge whether the stimulus pair was acoustically identical. The stimulus set was produced by reducing or increasing the closure duration between the successive moras in 5-ms steps (-95 to 240 ms). Results showed that the participants discriminated more precisely around closures duration ranging from approximately 150 to 180 ms (/bipa/: 180 ms, /guku/: 160 ms, /kuku/: 155 ms, /kuto/: 175 ms). Although the results indicate, to some degree, the possibility of perceptual categorization of the moraic obstruent, the high sensitivity peak in each stimulus type was slightly shifted from the perceptual boundaries in the identification task.

5pSC19. Intrinsic factors of releasing motions in an articulator: On assuming segmental input in speech-production models. Victor J. Boucher (Univ. of Montreal, C.P. 6128 succ. Centre-ville, Montreal, QC, Canada J3C 3J7)

Prominent models of speech production use serial input that is assumed to be commensurate with linguistic segments. The view is that such units underlie a serial activation of aperture motions such as closing and opening motions of the lips in articulating a bilabial stop. This is incompatible with conventional EMG observations showing a single burst of activity of labial adductors at the onset of a close-open cycle. The present study examines the spring-like effects of bilabial compression and pressure on labial opening (release) following a relaxation of the orbicularis oris muscle. Using reiterative series [papapapa] produced at increasing intensities, the range and velocity of opening motions of the lower-lip were correlated with lip compression and oral-pressure. The results for three speakers show that pressure and compression are correlated and that these factors account for 45% to 66% of the variance in velocity of lower-lip opening, and for 47% to 73% of the variance in the range of lower-lip opening. These results complement earlier findings of Abbs and Eilenberg (1976) showing the intrinsic effects of muscle elasticity on opening motions of the lips. Close-open cycles in articulators may not reflect segment-by-segment serial activation.

5pSC20. Kinematic properties of stop and fricative production. Anders Lofqvist (Haskins Labs., 270 Crown St., New Haven, CT 06511, lofquist@haskins.yale.edu)

This study examined movement kinematics in stop and fricative consonants, looking for potential evidence of different movement characteristics related to precision requirements. It has often been claimed that fricatives require more precision in their production than stops due to the management of airflow. However, the empirical evidence for this claim is

very limited and mostly based on different developmental trends. Tongue movements were recorded using a magnetometer system. Speed and movement paths were used to examine movement kinematics, including different measures of stiffness. Preliminary results do not show any consistent pattern across subjects, although some subjects did have systematic differences in movement stiffness with stop consonants having stiffer movements than fricatives. These results most likely reflect that fact that speech is an over learned skill, so that any differences related to precision are very small, or nonexistent, in adults, and probably only apparent in the early stages of speech development. [Work supported by NIH.]

5pSC21. Haptic information enhances auditory speech perception.

Kristin M. Johannsdottir, Diana Gibrail, Gick Bryan, Ikegami Yoko (Dept. of Linguist., Univ. of British Columbia, Canada), and Jeff Muehlbauer (Univ. of British Columbia, Canada)

Studies of tactile enhancement effects on auditory speech perception [Reed *et al.*, JSHR, 1978, 1982, 1989] have traditionally used experienced, pre-trained subjects. These studies have left unanswered the basic questions of whether tactile enhancement of speech is a basic component of human speech perception or a learned association [Fowler, JPhon, 1986; Diehl and Kluender, *Ecol. Psych.*, 1989], and which aspects of the signal are enhanced through this modality. The present study focuses exclusively on tactile enhancement effects available to naive speakers. In a speech-masked environment, naive subjects were tasked with identifying consonants using the Tadoma method. Half of the stimuli were presented with tactile input, and half without. The results show that all subjects gained considerably from tactile information, although which part of the tactile signal was most helpful varied by subject. The features that contributed most to consonant identification were aspiration and voicing. [Work supported by NSERC.]

5pSC22. Categorization of spectrally complex non-invariant auditory stimuli in a computer game task.

Lori Holt and Travis Wade (Dept. Psych., Carnegie Mellon Univ., 5000 Forbes Ave., Pittsburgh, PA 15213, twade@andrew.cmu.edu)

This study examined perceptual learning of spectrally complex non-speech auditory categories in an interactive multi-modal training paradigm. Participants played a computer game in which they navigated through a three-dimensional space while responding to animated characters encountered along the way. Characters appearances in the game correlated with distinctive sound category distributions, exemplars of which repeated each time the characters were encountered. As the game progressed, the speed and difficulty of required tasks increased and characters became harder to identify visually, so quick identification of approaching characters by sound patterns was, although never required or encouraged, of gradually increasing benefit. After thirty minutes of play, participants performed a categorization task, matching sounds to characters. Despite not being informed of audio-visual correlations, participants exhibited reliable learning of these patterns at post-test. Categorization accuracy was related to several measures of game performance and category learning was sensitive to category distribution differences modeling acoustic structures of speech categories. Category knowledge resulting from the game was qualitatively different from that gained from an explicit unsupervised categorization task involving the same stimuli. Results are discussed with respect to information sources and mechanisms involved in acquiring complex, context-dependent auditory categories, including phonetic categories, and to multi-modal statistical learning.

5pSC23. Acoustic correlates of non-modal phonation in telephone speech. Tae-Jin Yoon (Dept. of Linguist., Univ. of Illinois, 4080 FLB, 707 S. Mathews Ave., MC-168, Urbana, IL 61801, tyoon@uiuc.edu), Jennifer Cole, Mark Hasegawa-Johnson, and Chilin Shih (Univ. of Illinois, Urbana, IL 61801)

Non-modal phonation conveys both linguistic and paralinguistic information, and is distinguished by acoustic source and filter features. Detecting non-modal phonation in speech requires reliable F_0 analysis, a problem for telephone-band speech, where F_0 analysis frequently fails. An approach is demonstrated to the detection of creaky phonation in telephone speech based on robust F_0 and spectral analysis. The F_0 analysis relies on an autocorrelation algorithm applied to the inverse-filtered speech signal and succeeds in regions of non-modal phonation where the non-filtered F_0 analysis typically fails. In addition to the extracted F_0 values, spectral amplitude is measured at the first two harmonics (H_1 , H_2) and the first three formants (A_1 , A_2 , A_3). F_0 and spectral tilt are measured from 300 samples of modal and creaky voice vowels, selected from Switchboard telephone speech using auditory and visual criteria. Results show successful F_0 detection in creaky voice regions, with distinctive low F_0 , and statistically significant differences between modal and creaky voice in measures of spectral amplitude, especially for measures based on H_1 . Our current work develops methods for the automatic detection of creaky voicing in spontaneous speech based on the analysis technique shown here. [Work supported by NSF.]

5pSC24. A consonant-vowel-consonant display as a computer-based articulation training aid for the hearing impaired.

Fansheng Meng, Eugen Rodel, and Stephen Zahorian (Dept. of Elec. and Comput. Eng., Old Dominion Univ., Norfolk, VA 23529)

Computer-based visual speech training aids are potentially useful feedback tools for hearing impaired people. In this paper, a training aid for the articulation of short Consonant-Vowel-Consonant (CVC) words is presented using an integrated real-time display of phonetic content and loudness. Although not yet extensively tested with hearing-impaired listeners, real-time signal processing has been developed, flow-mode displays have been developed, and accuracy tests have been completed with a large database of CVC short words. A combination of Hidden Markov Models and time-delay neural networks are used as the primary method for the acoustic-phonetic transformations used in the display. [Work partially supported by NSF grant BES-9977260.]

5pSC25. Stop consonant perception in silent-center syllables.

Mark DeRuiter (Univ. of Minnesota, 164 Pillsbury Dr. S.E., 115 Shevlin Hall, Minneapolis, MN 55455), Virginia S. Ramachandran, and Sheryl J. Rosen (Wayne State Univ., Detroit, MI 48202)

Silent-center (SC) syllables are consonant-vowel-consonant syllables with the steady-state vowel portion excised from them. These syllables are appealing for research of stop-consonants because they retain the complexity, brevity, and rapid changes inherent in formant transitions while presumably eliminating temporal masking by the vowel. However, questions exist as to whether or not SC syllables are processed in the same manner as their full-vowel (FV) counterparts (i.e. are they processed as speech units or as sounds?). Data is reviewed from a series of experiments which examined listeners discrimination, labeling, and response time for synthesized consonant-vowel-consonant syllables in FV and SC conditions. Results from 3 experiments with typical listeners reveal that: (1) discrimination is significantly better in the SC condition; (2) consonant labeling on a /bab/ to /dad/ continuum is poorer in the SC condition and is significantly poorer at the category boundary; (3) discrimination response-time (RT) is significantly shorter in the SC condition. Labeling and discrimination results reveal that listeners processed the stop-consonants in these SC syllables in a less-categorical manner than in the FV syllables. Taken together with significantly different response-times, these results may indicate that listeners utilize a different mode of processing for SC syllables.

5pSC26. Voice *F0* responses elicited by perturbations in pitch of auditory feedback during English speech and sustained vowels. Jay J. Bauer (Dept. of Commun. Sci. & Disord., Univ. of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, WI 53201, jbauer@uwm.edu) and Charles R. Larson (Northwestern Univ., Evanston, IL 60208)

Twenty-one native speakers of American English were asked to repeatedly inflect the pitch of either the first or second syllable of an English speech phrase or produce non-word sustained vowels while listening to amplified auditory feedback. Brief upward and downward perturbations in pitch of auditory feedback were introduced in real time shortly after vocal onset. Resultant changes in voice *F0* due to perturbations in pitch were compared across vocal conditions and perturbation direction. Data indicated that auditory feedback was used in real-time to stabilize voice *F0* during both normal English speech and sustained vowels. Incomplete compensatory corrections in voice *F0* due to the perturbations in pitch of auditory feedback were prevalent during both dynamic speech (gain: ~12.7%) and static sustained vowel phonations (gain: ~19.3%), and appear to be regulated by similar corrective mechanisms. However during dynamic speech, voice *F0* was modulated across syllable boundaries in a task-dependent manner according to syllable pitch inflection. Thus, voice *F0* responses appear to help maintain the underlying suprasegmental meaning of a speech phrase by delaying the onset of the voice *F0* response to coincide with the pitch-inflected syllable. [Work supported by NIH Grant No. F31 DC005874.]

5pSC27. The perception of speech rhythm: An investigation of inter and intra rhythmic class variability using delexicalized stimuli. Volker Dellwo (Dept. of Phonet. and Linguist., University College London, 4 Stephenson Way, London NW1 2HE, UK) and Petra Wagner (Universitaet Bonn, D-53115 Bonn, Germany)

It has been demonstrated that stress- and syllable-timed languages are differentiable in terms of parameters based on vocalic and consonantal interval durations, and that stress-timed languages are more variable in terms of these parameters. There is also significant between- and within-subject variability on these parameters, but it is unclear whether this production variability is perceptually salient in terms of rhythm. This issue was investigated in a series of experiments using delexicalized stimuli, in which the vocalic intervals were substituted by sinusoidal tones and the consonantal intervals by white noise. Subjects were asked to classify the stimuli according to the regularity of the tone sequences. The results thus far indicate that intra rhythm class variation, as measured by consonantal and vocalic interval duration variation, can be well perceived.

5pSC28. Palato-lingual contact patterns during Korean obstruent production; lax, aspirated and forced. Hirohide Yoshioka and Kim Soone (Inst. of Human Sci., Univ. of Tsukuba, Japan)

In Korean obstruents, there is a three-way distinction in manner of articulation; lax, aspirated and forced. The present study is to investigate the temporal and spatial details of tongue contacts with the hard palate during these three types of consonant production. The subject, a 30-year-old female of the Seoul dialect was instructed to read a set of various words beginning with these consonants, embedded in a frame sentence i-geo-/Cal/-ita (C=consonant). The palato-lingual contact patterns were recorded using dynamic electro-palatography. The results show that the area and duration of the palato-lingual contact are clearly different among these three types of consonants. As for stops, the palato-lingual contact is wider and longer in the forced stop production. In the affricate production, the patterns of the maximum contact area and the duration are as follows: wide and long in the forced type, middle in the aspirated, and narrow and short in the lax. These data strongly suggest that the temporal and spatial details of the palato-lingual contact may play an important role in manifesting the Korean three-way distinction, together with other articulatory adjustments, such as laryngeal and respiratory controls.

5pSC29. Vowel production and perception in French blind and sighted adults. Sophie Dupont and Lucie Menard (Universite du Province Quebec a Montreal, Departement de linguistique, CP 8888, succ. Centre-Ville, Montreal, Canada H3C 3P8, menard.lucie@uqam.ca)

It is well known that visual cues play an important role in speech perception and production. Early in life, blind speakers, who do not have access to visual information related to lip and jaw movements, show a delay in the acquisition of phonological contrasts based on these visible articulatory features [A. E. Mills, *Hearing by eye: The psychology of lip-reading*, pp. 145–161 (1987)]. It has also been claimed that blind speakers have better auditory discrimination abilities than sighted speakers. The goal of this study is to describe the production-perception relationships involved in French vowels for blind and sighted speakers. Six blind adults and six sighted adults served as subjects. The auditory abilities of each subject were evaluated by auditory discrimination tests (AXB). At the production level, ten repetitions of the ten French oral vowels were recorded. Formant values and fundamental frequency values were extracted from the acoustic signal. Measures of contrasts (Euclidean distances) and dispersion (standard deviations) were computed and compared for each feature (height, place, roundedness) and group (blind, sighted). Regression analyses between discrimination scores and produced contrasts were carried out. Despite between-speaker variability, results show an effect of speakers group (blind versus sighted) on the produced contrast distances.

5pSC30. The role of intonational phrasing in the elicitation of speech errors. Marianne Pouplier (TAAL, Univ. of Edinburgh, UK and Haskins Labs, New Haven, CT, pouplier@ling.ed.ac.uk), Mark Tiede (Haskins Labs, New Haven, CT), Stefanie Shattuck-Hufnagel (MIT R.L.E., Cambridge, MA), Man Gao, and Louis Goldstein (Yale Univ., New Haven, CT)

Recent work has proposed that some speech errors arise from an underlying coordination process in speech production in which gestures can assume grammatically illegal but dynamically stable coordination modes. The general premise is that shared gestural structure within a prosodic domain sets up the conditions under which stable coordination modes spontaneously emerge and produce errors. If certain gestures (e.g., coda consonants) recur every word but others recur less frequently (e.g., alternating initial consonants), the lower frequency gestures will increase in occurrence, resulting in gestural intrusion errors. This suggests that manipulation of the rhythmic domain through intonational phrasing should determine error patterns differentially. Since the gestural coupling which lies at the heart of the rhythmic synchronization approach takes time to establish, it is predicted that continuous repetitions of e.g. “cop tap tube cub” will lead to a maximum buildup of coupling strength and thus the most errors. If however phrasal grouping is imposed, as in “cop tap, tube cub,” fewer errors are hypothesized to occur. We test these predictions by having subjects repeat the same word string in different prosodic grouping conditions. Our initial results confirm that the more alternating consonants an intonational phrase contains, the more errors occur.

5pSC31. Effect of aspiration noise and spectral slope on perceived breathiness in vowels. Rahul Shrivastav and Maria Pinero (Dept. of Comm. Sci. & Disord., Univ. of Florida, 336 Dauer Hall, Gainesville, FL 32611)

Breathy voice quality is frequently encountered in both normal and dysphonic voices. Breathiness has been associated with an increase in the intensity of aspiration noise as well as changes in spectral slope [Childers and Ahn, 1995; Fischer-Jorgenson, 1967; Huffman, 1987; Klatt and Klatt, 1990]. Shrivastav and Sapienza (2003) found that subjective ratings of breathiness obtained from a group of listeners were highly correlated with certain measures calculated from the auditory spectrum of the vocal acoustic signal. One of these measures, the partial loudness of the harmonic energy, was related to both the aspiration noise and the spectral slope of the voices. In the present experiment, 10 young adult listeners with normal

hearing were presented with voices that varied systematically in terms of their aspiration noise and spectral slope. The stimuli, five male and five female voices, were generated using the Klatt synthesizer and were modeled after naturally occurring voices. Listeners rated breathiness for each of these stimuli using a 7-point rating scale. Results show the relative contribution of spectral slope and aspiration noise to the perception of breathiness.

5pSC32. Intensity variation in vowels across acoustic and auditory spectra. Ewa Jacewicz and Robert Allen Fox (Dept. of Speech and Hearing Sci., The Ohio State Univ., Columbus, OH 43210, jacewicz.1@osu.edu)

Changes to vowel intensity are affected primarily by an increased physiological effort of the speaker. A second source of intensity variation comes from immediate consonant environment of the vowel as a result of coarticulation [House & Fairbanks, *J. Acoust. Soc. Am.* **22**, 105–113 (1953)]. Variation in intensity distribution across the vowel spectrum was measured for eight American English vowels in ten symmetrical CVC environments (consonants included both stops and fricatives) at four different locations (at the vowels rms peak and at points corresponding to 20%, 50%, and 80% of the vowels duration). Three types of spectral intensity distribution were examined: (1) the relative amplitudes of formants $F1$ - $F4$; (2) the summed intensity in four contiguous frequency bands of 0–0.5, 0.5–1.0, 1.0–2.0 and 2.0–4.0 kHz [Sluijter and Van Heuven, *J. Acoust. Soc. Am.* **100**, 2471–2485 (1996)]; and (3) the intensity distribution following three stages of auditory processing (using an auditory filter bank consisting of 33 filters, equal loudness pre-emphasis, and intensity-loudness compression). The nature and size of the effects of spectral intensity variation as a function of consonant environment in the acoustic spectrum (formants) and in auditory pre-processing of the acoustic spectrum will be compared. [Work supported by NIDCD R03 DC005560-01.]

5pSC33. Voice onset time is shorter in high-frequency words. Mark VanDam and Robert Port (Dept. of Linguist., Indiana Univ, Bloomington, IN 47405)

Frequency of occurrence is known to have many effects on speech production [see J. Bybee, *Phonology and Language Use* (Cambridge, 2001)] including vowel quality, overall duration, rate of deletion, assimilation, coarticulation, etc. The current work addresses voice-onset time (VOT) in words with differing lexical frequency estimates from published materials and addresses whether words in a list exhibit similar effects to words in sentential context. Four talkers produced 20 low frequency words and 10 high frequency words four times each in isolation and again in non-idiomatic, sentential context. VOT was measured in monosyllabic content words with initial /t/. Results show that frequent words (e.g., *talk*, *table*) have a mean VOT roughly 10 ms shorter than less frequent words (e.g., *tal*, *taint*) ($p < 0.01$). The effect was significantly stronger for words in a sentence (e.g., *He will talk to his supervisor*) than in a list. These findings are consistent with linguistic theories that propose detailed auditory representations for words and finely controlled productions, but are in conflict with traditional, abstract phonological representations.

5pSC34. Tension asymmetries in a finite element model of the vocal folds. Greg S. Davidson (Univ. of Chicago, Chicago, IL 60637) and Fariborz Alipour (Univ. of Iowa, Iowa City, IA 52242)

Tension asymmetries in a finite element model of the vocal folds were examined as a function of lung pressure. The vocal fold model had an asymmetry in the tension ratio ranging from 1.1 to 2.0, corresponding to a 10%–100% increase in tension of the left fold compared to the right, and

lung pressure was increased over the range 5–40 cm H₂O. For tension ratios greater than 1.6, oscillation was not supported over most of this range. For tension ratios less than 1.6, the fundamental frequency was found to increase up to approximately 22 cm H₂O. Mean vocal fold contact area showed a broad maximum over the range 8–13 cm H₂O and decreased between 13 and 22–25 cm H₂O. Open quotient decreased between 7.5 and 13–15 cm H₂O and increased afterwards to reach a maximum value between 20 and 23 cm H₂O. For these three parameters, the behavior above the upper thresholds was tension ratio dependent. Mean glottal area linearly increased with lung pressure with differing slopes over different pressure ranges. Taken together, the four parameters provide evidence for different vibratory domains. [Work supported by NIDCD grant No. DC03566.]

5pSC35. Not just any contour: Intonation marking semantic domain narrowing. Karsten A. Koch (Univ. of British Columbia, Vancouver, BC, Canada)

Not just any statements (*Natalie won't drink just ANY wine from France*) carry an intonational contour. Previous studies have characterized the contour as fall-rise [D. Robert Ladd, *The structure of intonational meaning* (1980)], but more intricate acoustic analysis that considers the semantics of the construction has not been performed. Preliminary data shows that for speakers of Canadian English, the contour contains a rise-fall and stress accent on *any*, followed by a rise on the sentence-final syllable. If the stressed rise-fall on *any* indicates a topic accent, then the meaning of the construction can be calculated semantically: the presence of the topic accent indicates a disputable topic (in this case, Which wines will Natalie drink?) [following Daniel Buring, *Linguistics and Philosophy* **20**, 175–194 (1997) for German]. This operation narrows the domain of *any wine* from all wine from France to some smaller amount perhaps Natalie only drinks expensive cabernets. Furthermore, the same intonational contour occurs in other quantificational statements, like *ALL gamblers aren't addicted*. Here, the disputable topic is: How many gamblers are addicted? Again, the intonation serves to narrow the domain to a subset of all the gamblers. Thus, documenting this intonational contour identifies a more general semantic process.

5pSC36. The role of stress accent in the understanding of sentences in noise. Pierre Divenyi and Alex Brandmeyer (Speech and Hearing Res., VA Medical Ctr., 150 Muir Rd., Martinez, CA 94553)

When a target sentence is presented in noise, the stressed syllables therein will constitute epochs with a high likelihood of S/N increase. Since in spontaneous speech stressed syllables have significantly higher information content than unstressed syllables [S. Greenberg *et al.*, in *Proc. 2d Intern. Conf. Human Lang. Technol. Res.* (2002), pp. 36–43], it follows that masking the stressed syllables will negatively impact the perception of neighboring unstressed syllables. This hypothesis was tested by varying the S/N of only the stressed syllables in sentences embedded in otherwise unmodulated speech-spectrum noise. Results suggest that, compared to conditions with absent or decreased-S/N stressed syllables, increasing the S/N during stressed syllables increases intelligibility of the unstressed syllable or syllables that follow. Since the unstressed syllables are rarely masked in the classic, energetic sense, the phenomenon reported constitutes an instance of informational masking. The masking can affect perception at the phonemic as well as on higher linguistic levels, depending on the information provided by sentence context.

Session 5pUW

Underwater Acoustics: Autonomous Underwater Vehicles (AUV's), Sensors, and Arrays

Brian T. Hefner, Chair

Applied Physics Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698

Contributed Papers

1:30

5pUW1. An autonomous underwater vehicle towed hydrophone array for ocean acoustic measurements. Jason D. Holmes, William M. Carey (Boston Univ., Boston, MA 02215), James F. Lynch, Arthur E. Newhall, and Amy Kukulya (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

An autonomous underwater vehicle (AUV) with a towed hydrophone array (THA) can provide the capability of mobile-single-ship operation for both short-range single path and long range synthetic aperture experiments. A low noise towed array for an AUV (REMUS) has been developed to demonstrate the feasibility and utility of such measurements. Previous measurements of AUV radiated noise indicated levels that would limit measurements by hull arrays providing a rationale for a THA. A small-diameter fluid-filled prototype hydrophone array was developed and tested to ensure adequate sensitivity and system noise levels. The digital recording system (DRS) consisted of mini-disc recorders with a band width of 20 kHz housed in a tube attached to the AUV. This combined system (REMUS, DRS, and THA) was used to conduct a proof of concept test in Dodge Pond. This paper presents the results that show, in the Sea State 0 noise field of Dodge Pond, array system self noise was less than the ambient and vehicle noise was manageable. The test included three pingers and a beacon sound source. The acoustic results were at a high signal-to-noise ratio and show the array tow stability and coherent processing capability. [Work supported by Boston University, Woods Hole, and ONR.]

1:45

5pUW2. Response pattern control in spite of failed array elements. Henry Cox and Hung Lai (Lockheed Martin Orincon, 4350 North Fairfax Dr., Ste. 470, Arlington, VA 22203)

Large arrays of many elements usually are shaded to achieve low sidelobes. However, in practice these arrays typically operate with some element failures that significantly degrade sidelobe performance. First, the effects of different numbers of failed elements on sidelobe levels are analyzed and illustrated with simulations. Statistics and bounds for sidelobe levels with failed elements are presented. The inability and limitations of conventional shading modifications to compensate for failed elements is discussed and illustrated. Missing elements give rise to large inter-element spacings that in turn result in grating effects that cannot be compensated for using modified shading weights. An adaptive compensation approach for failed elements is then presented. This involves narrowband interpolation from nearby elements to estimate the missing data prior to applying the original shading weights. Failures of both isolated and neighboring elements are considered. When there is a strong source present it is relatively easy to estimate its value on missing elements. But this is exactly the case of interest, when sidelobe levels are important. The adaptive interpolation approach provides good bearing response patterns so that strong signals are well isolated in bearing. Results are illustrated with large array simulations with different failure patterns and multiple strong sources.

2:00

5pUW3. Underwater acoustic measurements with the Liberdade/X-Ray flying wing glider. Gerald L. D'Spain, Scott A. Jenkins, Richard Zimmerman (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 93940-0701), James C. Luby (Univ. of Washington, Seattle, WA 98105), and Aaron M. Thode (Scripps Inst. of Oceanogr., La Jolla, CA 93940-0701)

An underwater glider based on a flying wing design (Jenkins *et al.*, 2003) presently is under development by the Marine Physical Laboratory, Scripps Institution of Oceanography and the Applied Physics Laboratory, University of Washington. This design maximizes the horizontal distance between changes in buoyancy to minimize mechanical power consumed in horizontal transport. The prototype wing has a 6.1 m wing span and is 20 times larger by volume than existing gliders. Initial at-sea tests indicate that the lift-to-drag ratio is 17/1 at a horizontal speed of about 1.8 m/s for a 38-liter buoyancy engine. Beamforming results using recordings of the radiated noise from the deployment ship by two hydrophones mounted on the wing verify aspects of the prototype wing flight characteristics. The payload on the new glider will include a low-power, 32-element hydrophone array placed along the leading edge of the wing for large physical aperture at midfrequencies (above 1 kHz) and a 4-component vector sensor. Data previously collected by these types of arrays illustrate the performance of narrow-band detection and localization algorithms. Flight behaviors are being developed to maximize the arrays' detection and localization capabilities. [Work sponsored by the Office of Naval Research.]

2:15

5pUW4. Adaptive rapid environmental assessment simulation framework. Ding Wang and Henrik Schmidt (Acoust. Lab, Ocean Eng. Dept., MIT, 77 Mass Ave 5-435, Cambridge, MA 02139, prolog@mit.edu)

The coastal environment is characterized by variability on small spatial scales and short temporal scales. These environmental uncertainties translate—in a highly non-linear fashion—into uncertain acoustic propagation properties, often severely affecting the sonar performance prediction capabilities. Conventional oceanographic measurement systems cannot capture these environmental uncertainties due to their limited predictability and the lack of mobility of the traditional measurement platforms. Taking advantage of the mobility of modern autonomous underwater vehicle technology, the Adaptive Rapid Environmental Assessment (AREA) concept has been developed to optimally use the available resources to capture the acoustic uncertainty by adaptive and rapid *in situ* measurement of the environmental properties most significant to the actual sonar system. The ocean area of interest is usually fairly large, and the *in situ* measurement resources are limited, and the adaptive sampling strategy in AREA can therefore make a significant difference in capturing and mitigating the sonar performance uncertainties. To determine an optimal or sub-optimal sampling strategy and test the optimization effect before doing costly on-site experiments, an Adaptive Rapid Environmental Assessment Simulation Framework (AREASF) has been developed, based on state-of-the-art forecasting frameworks and acoustic propagation models. The simulation framework has been applied to investigate the performance of AREA under different environmental conditions, and based on these

results, the feasibility of performing real-time adaptive environmental assessment using AREA under realistic ocean conditions will be discussed. [Work supported by ONR.]

2:30

5pUW5. Optimizing multistatic sonobuoy placement. Donald R. DelBalzo, Erik R. Rike, and David N. McNeal (Neptune Sci. Div. of Planning Systems, Inc., 40201 Hwy 190 E, Slidell, LA 70461, delbalzo@neptunesci.com)

Sonobuoy patterns for monostatic sensors were developed during the Cold War for deep, uniform underwater environments, where a simple median detection range defined a fixed inter-buoy spacing (usually along staggered lines). Oceanographic and acoustic conditions in littoral environments are so complex and dynamic that spatial and temporal variability of low-frequency signal and noise fields destroys the basic homogeneous assumption associated with standard tactical search concepts. Genetic Algorithms (GAs) have been applied to this problem to produce near-optimal, non-standard search tracks for monostatic mobile sensors that maximize probability of detection in such inhomogeneous environments. The present work describes a new capability, SCOUT (Sensor Coordination for Optimal Utilization and Tactics), to simulate multistatic distributed-sensor geometries and to optimize the locations of multistatic active sonobuoys in a complex, littoral environment. This presentation reviews the GA approach, discusses the new chromosome structure, and introduces a new target-centric geometry. The results show that (a) standard patterns are not optimal even for a homogeneous environment, (b) small distributed sensor clusters are preferred, and (c) standard patterns are grossly ineffective in inhomogeneous environments where 20% improvements in detection are achieved with SCOUT. [Work supported by NAVAIR.]

2:45

5pUW6. Acoustic particle velocity and intensity calculations from tri-axial pressure gradient measurements. Melanie E. Austin and Alex O. MacGillivray (JASCO Res. Ltd., 2101-4464 Markham St., Victoria, BC, Canada V8Z 7X8, melanie@jasco.com)

In July 2004 Fisheries and Oceans Canada supported a study to investigate effects of seismic airgun signals on hearing organs of freshwater fish

in the Mackenzie River at Inuvik, NWT Canada. The study required particle velocity measurements for correlation with observed biological effects. JASCO Research built a pressure gradient measurement apparatus consisting of four hydrophones mounted at the vertices of a triangular-pyramid frame. The system was used to measure differential pressure from the airgun events simultaneously in three perpendicular axial directions. An attached depth-compass sensor monitored the depth and orientation of the system. Hydrophone separations were chosen to be small relative to the acoustic wavelength so that measured differential pressures correctly approximated the pressure gradients along each axis. Particle accelerations were computed directly from pressure gradients following Euler's linearized momentum equation, and particle velocities were computed by integrating particle accelerations. Acoustic intensity was computed from the product of acoustic pressure and particle velocity. The hydrophone precision imposed a limit on accuracy of particle velocity measurements at low frequencies. Likewise the fixed hydrophone spacings defined an upper frequency limit for applicability of this method.

3:00

5pUW7. Macroscopic sonic crystals in a wave guide. Dalcio K. Dacol, Gregory J. Orris, and David C. Calvo (Acoust. Div. Naval Res. Lab., Washington, DC 20375-5350)

Periodical arrays of long, parallel cylinders have been shown in the published literature to exhibit acoustical properties which are analogous to the electronic properties of certain crystalline solids. Thus there are frequency bands for which no acoustic propagation through the array is allowed, with those band gaps being directionally dependent. The properties of such arrays in wave guides are examined in this work. Particular emphasis is put in investigating the properties of those arrays in wave guides that are idealized models of shallow water oceanic wave guides. The effects of those arrays on the propagation of normal modes is investigated in detail. The possibility of using such arrays in the construction of large scale underwater devices such as filters and lenses is also discussed. [Work supported by ONR.]